Reasoning about morality

Verhoeven, Willem Adriaan

DOI:
10.33612/diss.135594000

Reasoning about Morality

The Empirical Turn in Early-Modern Moral Philosophy from Hobbes to Rousseau

Willem Verhoeven

Reasoning about Morality: The Empirical Turn in Early-Modern Moral Philosophy from Hobbes to Rousseau
Thesis, University of Groningen, Groningen, The Netherlands

Cover image Bart Geelen
Printing Proefschriftmaken.nl
ISBN 978-94-034-2882-6

rijksuniversiteit
groningen

Reasoning about Morality

The Empirical Turn in Early-Modern Moral Philosophy from Hobbes to Rousseau

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 3 september 2020 om 11:00

door

Willem Adriaan Verhoeven

geboren op 14 december 1985
te Amersfoort

Promotores
Prof. dr. L.W. Nauta
Prof. dr. A.M.R. de Dijn

Beoordelingscommissie
Prof. dr. M. Lenz
Prof. dr. A. Lifschitz
Prof. dr. J.A. van Ruler

Abstract

This thesis investigates the accounts of human understanding proposed by Thomas Hobbes, John Locke, the Abbé de Condillac, Claude-Adrien Helvétius and Jean-Jacques Rousseau, and explores the consequences of their epistemologies for their theories of morality and natural law. In response to the ethical question how humans may determine the morality of their conduct by means of reason, the five thinkers we will investigate draw normative conclusions from a descriptive account of human nature. We will thus see that their descriptive account of the role of both the passions and reason in the determination of our actions provides the basis for a normative account of moral judgment. For the five thinkers we will investigate, reason is thus not merely the ability to produce knowledge by comparing disparate ideas, but also a faculty involved in practical deliberation, which not only allows us to determine whether our actions are prudent, but ultimately also whether they are moral. Accordingly, their account of human understanding not only constitutes their epistemology, but is also an integral part of their moral philosophy. As the epistemology of these five thinkers is clearly empirical in outlook, this thesis will document the impact of the empirical turn on early-modern moral philosophy. This enquiry will provide grounds to reconsider some prevailing scholarly opinions on the relation between reason and the passions in early-modern moral philosophy. In addition, our discussion will aim to clarify the relation between facts and values in the theories of several early-modern thinkers predating Hume. Finally, this thesis will trace the development of early-modern consequentialist accounts of normative judgment towards more explicit formulations of utilitarianism, and discuss the objections of one of the most important early-modern opponents of consequentialism – Jean-Jacques Rousseau.

Table of Contents

Introduction

In this thesis, we will investigate the accounts of human understanding proposed by Thomas Hobbes, John Locke, the Abbé de Condillac, Claude-Adrien Helvétius and Jean-Jacques Rousseau, and explore the consequences of their epistemologies for their theories of morality and natural law. These five authors wrote in widely different contexts and over the course of more than a century. Nonetheless, the fact that these thinkers shared a significant number of philosophical preconceptions and preoccupations justifies their selection for our discussion. Thus all five thinkers advance an empiricist account of human understanding, according to which reasoning is nothing but the comparison or relating of ideas acquired through experience and organised with the help of language.[1] Furthermore, these philosophers accept a hedonist account of human motivation, maintaining that all human thoughts and actions are occasioned by a desire for pleasure or aversion of pain. With the possible exception of Locke and Rousseau, these thinkers also conceive of morality as a prudential guideline that indicates the means to felicity, conceived as the durable satisfaction of the passions. Finally, all five philosophers maintain that this moral guideline may be deduced by means of reasoning.

Yet by their own account of the faculty of reason, it is by no means obvious how these philosophers think that reasoning could provide human beings with an understanding of morality or natural law. All five thinkers describe the process of reasoning as the comparison of ideas derived from experience and structured with the help of language. This account of reasoning describes relatively clearly how human beings use reason to acquire knowledge of empirical reality. Yet as we will see, this same account does not yet provide a method by which reason can produce normative conclusions. As the five thinkers featured in this thesis nonetheless claim that morality and natural law are

[1] In the case of Hobbes and Rousseau, their characterisation as empiricists may be disputed. Thus while Hobbes clearly develops an empiricist account of human understanding, certain elements of his scientific method, such as his insistence on the use of syllogistic logic and his conception of science as a unified body of knowledge, diverge from the empiricist tradition as it developed from Locke onwards. While Rousseau likewise adopts an empiricist account of human understanding, he also introduces some innate properties and faculties to his conception of human nature. Observing these caveats, our reading will nonetheless interpret both Hobbes and Rousseau as empiricist thinkers.

accessible by means of reason, we will try to reconstruct their account of normative reasoning. The starting point of our investigation will thus be the question how these philosophers think that human beings should use their reason in order to understand morality and natural law. The resulting discussion will then provide us with new perspectives on the ethical theories of these five thinkers, as well as on the history of early-modern moral philosophy as a whole.

Of course, scholars have previously discussed the theories of these five philosophers at length, as well as the history of early-modern moral philosophy in general, placing them in a variety of contexts. Thus many scholars have already described how seventeenth-century philosophers became particularly concerned with the question of the foundations of morality following the demise of the scholastic metaphysics that had buttressed previous theories of ethics.[2] Others, most notably Richard Tuck, have maintained that seventeenth-century moral philosophers including Grotius and Hobbes aimed primarily to counter the relativist arguments of ancient scepticism, which had experienced something of a revival in the period.[3] Alternatively, Stephen Gaukroger has placed early-modern ethics in the context of the scientific revolution, claiming that the transformation of the norms and methods of the natural sciences also shaped the moral philosophy of the era.[4] Furthermore, Tim Hochstrasser has shown that the modern perception of early-modern natural law has been shaped by the early-modern historiography of moral philosophy.[5] Finally, early-modern ethics also features prominently in the historiography of the Enlightenment, as intellectual

[2] Stephen Darwall, 'The Foundations of Morality', in: Donald Rutherford (ed.), *The Cambridge Companion to Early Modern Philosophy* (Cambridge 2006) 221-247, 223; Paulien Westerman, *The Disintegration of Natural Law Theory: Aquinas to Finnis* (Leiden 1998).

[3] Richard Tuck, 'The 'Modern' Theory of Natural Law', in: Anthony Pagden (ed.), *The Languages of Political Theory in Early-Modern Europe* (Cambridge 1987) 99-119. Knud Haakonssen, *Natural Law and Moral Philosophy: from Grotius to the Scottish Enlightenment* (Cambridge 1996) 25 agrees with this view. On scepticism and its relation to Hobbes's thought in particular see Gianni Pagannini, 'Hobbes and the French Skeptics', in: John Christian Laursen and idem, *Skepticism and Political Thought in the Seventeenth and Eighteenth Centuries* (Toronto 2015) 55-82 and Tuck, 'Optics and Sceptics: the Philosophical Foundations of Hobbes's Political Thought', in Edmund Leites (ed.), *Conscience and Casuistry in Early Modern Europe* (Cambridge 2002). This interpretation of Hobbes as motivated by scepticism is contested by Perez Zagorin, 'Hobbes's early philosophical development', *Journal of the History of Ideas* 54 (1993) 505-518; Zagorin, 'Hobbes without Grotius', *History of Political Thought* 21 (2000) 16-40.

[4] Stephen Gaukroger, *The Collapse of Mechanism and the Rise of Sensibility: Science and the Shaping of Modernity 1680-1760* (Oxford 2010).

[5] Tim Hochstrasser, *Natural Law Theories in the Early Enlightenment* (Cambridge 2000).

historians have documented how philosophical ideas inspired contemporary discussions on social and political reform.[6]

These are of course all valuable perspectives on the history of early-modern moral philosophy. In this thesis we will try to complement these scholarly achievements by analysing the moral philosophies of several important early-modern thinkers in the context of their theories of epistemology. Even if the relation between early-modern ethics and epistemology has not been entirely ignored in the literature, more recent scholars have generally had other preoccupations.[7] In fact, some prominent historians have explicitly questioned the continued relevance of epistemology to the interpretation of early-modern moral philosophy. Thus Knud Haakonssen laments that the historiography of early-modern philosophy has been distorted by an 'epistemological paradigm,' according to which all philosophical questions are seen as subsidiary to the question of the origin of knowledge.[8] Accordingly, Haakonssen dismisses the notion that early-modern epistemology also had consequences for morality as a Kantian anachronism.[9] A similar point has been made by J.B. Schneewind, who criticises the approach of scholars who would treat moral philosophy as merely

[6] Ernst Cassirer, *The Philosophy of the Enlightenment* (Princeton 2009 [1951]) is perhaps still the most illuminating study of the philosophical theories behind Enlightenment ideals. Anthony Pagden, *The Enlightenment and why it still Matters* (Princeton 2013) is a more recent synthesis of Enlightenment thought that also relates social and political debates to philosophical ideas. The most comprehensive recent study of the Enlightenment is presented by Jonathan Israel in his trilogy *Radical Enlightenment: Philosophy and the Making of Modernity 1650-1750* (Oxford 2001); *Enlightenment Contested: Philosophy, Modernity, and the Emancipation of Man 1670-1752* (Oxford 2006); *Democratic Enlightenment: Philosophy, Revolution, and Human Rights 1750-1790* (Oxford 2011). Yet while Israel is expansive in the selection of source material, his interpretation is rather sweeping, as he attempts to classify all Enlightenment thinkers into two opposing schools of thought based on their ontological preconceptions. Israel is therefore hardly interested in epistemology and its possible relevance to ethics. For a critical and insightful review of the shortcomings of Israel's approach, see Anthony la Vopa, 'A New Intellectual History? Jonathan Israel's Enlightenment', *The Historical Journal* 52 (2009) 717-738.

[7] It is for instance acknowledged by Cassirer, *The Philosophy of the Enlightenment* 108.

[8] Knud Haakonssen, 'The History of Eighteenth-Century Philosophy: History or Philosophy?', in: idem (ed.), *The Cambridge History of Eighteenth-Century Philosophy* (Cambridge 2006) I 3-25, 14.

[9] Haakonssen, 'Natural Law without Metaphysics: a Protestant Tradition', in: Ana Martinez Gonzalez (ed.), *Contemporary Perspectives on Natural Law* (London 2008) 67-85, 72-73. Haakonssen maintains that (moral) knowledge for Lockeans was not propositional. Yet this would seem to contradict with the insistence of Locke and Hobbes and others that morality would be demonstrated through reasoning with concepts/words, forming a law.

derivative of a thinker's ontology and epistemology.[10] Although Schneewind admits that epistemology and moral philosophy were not isolated fields of enquiry, he maintains that generally it was the moral views of thinkers that influenced their epistemology and not the other way around.[11]

Yet in our own discussion of five prominent early-modern philosophers, we will see that they did not apply any strict demarcation between moral philosophy and epistemology – a term that was in fact unknown at the time.[12] Thus Locke's *Essay concerning Human Understanding* is nowadays read as primarily a treatise on epistemology. Yet Locke likely wrote the *Essay* in response to debates on morality and revealed religion.[13] Indeed, a close inspection of the *Essay* reveals that while at first glance many of its chapters do not appear to be directly concerned with ethics, most of its theories do have important consequences for morality. Conversely, Hobbes's *Leviathan* is today rightly considered as a landmark in moral and political theory. Yet Hobbes does not commence his deduction of natural laws and exposition of the rights of citizens and sovereign before providing a theory of human understanding indicating the method by which humans should consult their reason in order to guide their conduct. Finally, we will see that Rousseau is another thinker whose account of human understanding is an integral part of his moral philosophy. While Rousseau is not primarily interested in epistemology in order to provide a theory of knowledge, his objective is to document how the development of human understanding has transformed man from a solitary but happy creature into a social yet miserable being.[14] As Rousseau considers the development of

[10] J.B. Schneewind, *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge 1998) 9-10.

[11] Schneewind, 'Voluntarism and the Foundations of Ethics', in: idem, *Essays on the History of Moral Philosophy* 202-221, 203: 'Philosophers did not simply apply to morality the theoretical views they developed in response to problems posed by scepticism and the new science. Rather, substantive moral beliefs imposed conditions on the theories of knowledge.'

[12] For brevity's sake, we will nonetheless employ the modern term 'epistemology' to refer to early-modern accounts of human understanding and associated theories of knowledge.

[13] G.A.J. Rogers, 'The Intellectual Setting and the Aims of the *Essay*', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* (Cambridge 2007) 7-32, 8.

[14] Throughout our discussion we will sometimes employ the term 'man' to refer to all human beings, both man and women, in an abstract sense, thereby following the practice of Hannah Dawson, who in her *Locke, Language and Early-Modern Philosophy* (Cambridge 2007) 1n1 explains that: 'Early-modern philosophers generally speak of 'men,' rather than 'men and women.' In order to avoid anachronism or exculpation, I tend to maintain this usage.'

morality as inextricable from the progress of human understanding, his writings provide an eighteenth-century example of a philosophy in which ethics and epistemology are interrelated rather than separate subjects. These observations then suggest that the endeavour of exploring the relations between early-modern ethics and epistemology is not as much based on anachronistic assumptions as the idea that these topics should be considered separate fields of inquiry.

For this reason, our own reading will take a more holistic approach, as we will interpret the moral philosophy of five early-modern thinkers on the premise that their views on morality are not isolated positions, but constitute a more or less coherent philosophy. This coherence is reflected in the methodology employed by the five thinkers we will investigate, as they all present a descriptive account of human nature as the foundation of both their theory of epistemology and moral philosophy. Like most other early-modern empiricists, these philosophers thereby formulate answers to questions about both knowledge and morality by investigating how human beings think and act in practice. Thus confronted with the question how humans may acquire knowledge, they first reconstruct how the contents of the mind is acquired through experience and structured with the help of language. Subsequently these thinkers describe meticulously how the process of reasoning may turn disparate ideas into knowledge. This descriptive account of human understanding then implies how human beings should conduct their reasoning in order to gain reliable knowledge of empirical reality.[15]

In response to the moral question how humans should conduct themselves towards each other, the five thinkers we will investigate also draw normative conclusions from a descriptive account of human nature – albeit in a slightly different manner. They begin this account of human nature by determining what motivates people, describing the process by which alternating desires and aversions provide the impetus for all our thoughts and actions. Yet most early-modern empiricists recognise that human beings are not merely automatons that blindly follow their desires. These philosophers describe how

[15] This approach common to most empiricist theories of deriving normative epistemological conclusions from a descriptive account of human understanding is dismissed by Richard Rorty, *Philosophy and the Mirror of Nature* (Princeton and Oxford 2009 [1979]) 139-148. For Rorty, the problem of this approach is not only that it is impossible to derive normative conclusions from a descriptive account of human understanding. According to Rorty, this account of human understanding is itself based on 'armchair reflection,' rather than empirical enquiry, thereby placing the whole of empirical science on foundations that have not been empirically verified.

humans may employ their reason to determine which passions are worthwhile pursuing, and which are better ignored. In other words, they provide an account of practical deliberation. This descriptive account of the role of both the passions and reason in the determination of our actions then provides the basis for a normative theory on how humans should use their reason to determine the morality of their conduct. For the five thinkers we will investigate, reason is thus not merely the ability to produce knowledge by comparing disparate ideas, but also a faculty involved in practical deliberation, which allows us to determine whether our actions are prudent. Yet as these thinkers, with the exception of Rousseau, consider morality as a prudential guideline, practical deliberation by means of reasoning is also at the basis of moral judgment. Accordingly, not only the epistemology but also the moral philosophy of these early-modern empiricists relies on an account of human understanding. Ultimately, this is the most important reason why this thesis will investigate the moral philosophy of five early-modern thinkers in the context of their theory of human understanding. And as the epistemology of these five thinkers is clearly empirical in outlook, we will thereby document the impact of the empirical turn on early-modern moral philosophy.

Main Themes and Research Questions

As we discuss the conceptions of human nature of Hobbes, Locke, Condillac, Helvétius, and Rousseau in order to discover how they think reasoning may allow us to understand morality, we will encounter a number of recurring topics and associated questions of interpretation. These issues are not equally relevant to all five thinkers, as they each have their own preoccupations. Nonetheless, we will find that there are at least five subjects that feature to some extent in all the chapters of this thesis. In the following, we will briefly outline these five recurring issues and formulate several questions that will inform our interpretation. The order in which these subjects and questions are displayed in this section roughly represents the structure of the chapters of our main discussion.

As we aim to interpret the moral philosophy of five thinkers in the context of their epistemology, our first task will be to provide an overview of their accounts of human understanding. We will then find that all the five philosophers agree that reasoning is nothing but the activity of comparing and relating ideas, which they consider as the cognitive units that provide the

contents of our thoughts. Furthermore, they all concur that the contents of these ideas may ultimately be traced back to our experience of empirical reality. While this preconception describes the origin of ideas of empirical objects, it does not yet explain the means by which human beings may acquire normative concepts like 'justice' or 'property,' as the contents of these notions is not simply experienced by means of the senses. Our investigation will therefore start by considering the question: by what means do the five philosophers envision that human beings acquire their ideas of morality? Our answer to this question will then reveal that many of these philosophers in fact think that moral concepts are a distinct type of idea separated from our ideas of empirical reality. Accordingly, we will also consider the question to what extent this dichotomy between moral concepts and ideas of empirical objects is relevant to our interpretation of their moral philosophy.

Subsequently, we will further investigate the epistemology of Hobbes, Locke, Condillac, Helvétius, and Rousseau in order to reconstruct their conception of normative reason. We will see that these philosophers all deny that reason is an autonomous faculty somehow innate to human nature that may deduce eternal truths independent from experience. They describe reason rather as simply the activity of comparing and relating ideas present in the mind. Yet we will see that while they agree on this presupposition, the five philosophers diverge in their account of the mechanics of reasoning. Thus some, such as Helvétius and arguably also Hobbes and Condillac, portray the process of reasoning as nothing other than the self-sustained interaction of ideas present in the mind. Others, including Locke and Rousseau, assume that the mind contains a distinct faculty of judgment that reasons by actively comparing ideas. Accordingly, some scholars have ascribed to these thinkers the theory that this faculty of judgment is also responsible for normative reasoning, leading to an interpretation according to which the ability for moral reasoning is innate to human nature. In our own investigation, we will therefore not only consider the question: to what extent do the five philosophers think that reason depends on abilities innate to human understanding? We will also inquire if they indeed claim that normative judgment is an ability innate to human nature as well.

Another recurring issue in our discussion will be the role of language in normative judgment. All five philosophers consider language not merely as a method of communication, but also maintain that the use of words is indispensible to the composition of complex ideas, and thereby to the full

development of human understanding. In fact, several thinkers claim that the ability for reasoning is only made possible by the use of language. Yet this theory that the use of words facilitates reasoning may also suggest that language could influence or even determine our thoughts. We will see that the philosophers following Locke were particularly sensitive to the fact that the signification of words is circumscribed by conventions upheld by the tacit consent of the community of speakers. In order to be understood, individuals are thus required to conform the signification of terms to linguistic conventions. In combination with the theory that the composition of complex ideas is only made possible by the use of words, this position suggests that individuals accommodate the composition of their ideas to established linguistic conventions. Accordingly, the conventions of language would also direct normative reasoning. As there have been scholarly interpretations that ascribe this conclusion to several of the philosophers discussed in this thesis, we will reconstruct their precise views on the role of language in cognition. Subsequently, we will pose the question: do any of these five philosophers believe that as words fulfil a crucial cognitive function, this also entails that the meaning of these words determines thought processes including normative judgment?

A fourth recurring issue in our discussion will be the role of the passions in the moral epistemology of Locke, Hobbes, Condillac, Helvétius, and Rousseau. We will see that these thinkers agree that all human thoughts and actions are ultimately motivated by a desire for pleasure or an aversion of pain. Furthermore, they equally reject the Aristotelian notion that human nature suggests a specific purpose that all persons should strive to fulfil in order to achieve moral and achieve happiness. In other words, these five philosophers dismiss the notion of a *summum bonum.* Alternatively, they maintain that human beings find happiness simply by fulfilling their passions, and accordingly have no other clearly defined purpose in life. Furthermore, the five thinkers examined in this thesis maintain that the passions also determine subjective normative judgment, as individuals consider the objects of their desire as *good,* while they value actions with painful consequences as *evil.* Altogether, these positions raise at least two questions that will be addressed in this thesis. First of all, if the passions provide the ultimate motivation for all thoughts and actions, what is the role of reason in practical deliberation? Secondly, if there is no *summum bonum*, as the object of the passions and thereby subjective normative judgment diverges

considerably among individuals, how could we devise a moral guideline that nonetheless directs all individuals towards felicity?

Following our discussion of these particular elements of the conceptions of human nature of the five philosophers, we will finally outline how these components fit together to form a coherent account of moral epistemology. We will then be in the position to answer our main question by reconstructing the method by which they think that reason may provide insight into morality and natural law. Yet on the basis of our investigation of the moral epistemology of these thinkers, we will also be able to contribute to the interpretation of their moral philosophy as a whole. Thus in several of the chapters of this thesis, we will conclude our discussion by considering the following two questions: in view of the conclusions of our investigation into their moral epistemology, what kind of moral theory did these philosophers propose? And: does our reconstruction of their accounts of moral epistemology suggest that any common interpretations of the moral philosophy of Hobbes, Locke, Condillac, Helvétius, and Rousseau must be amended?

Approach and Methodology

The selection of Hobbes, Locke, Condillac, Helvétius, and Rousseau for our discussion may at first appear somewhat arbitrary, as these thinkers published their primary writings over the course of a century in widely different social and political contexts. In addition, this selection ignores a great number of thinkers whose theories might also deserve attention in a discussion of early-modern moral philosophy.[16] Admittedly, the choice to include these particular five thinkers in in our investigation was not entirely premeditated, as expanding interests led research for this thesis down unforeseen paths. In accordance with our expansive selection of authors, each chapter in this thesis will present a largely self-contained discussion. Of course, these discussions will often provide references to other chapters. Furthermore, as we have outlined above, there are several recurring issues that feature in all our subsequent chapters. Nonetheless,

[16] Probably the most obvious candidate to include in our discussion would have been David Hume. Furthermore, following appraisal of Schneewind, 'Pufendorf's Place in the History of Ethics', *Synthese* 72 (1987) 123-155, the inclusion of Samuel Pufendorf would also have been a valuable addition. Yet as the inclusion of these two thinkers would widen the scope of this already expansive discussion to include the contexts of both the German and Scottish Enlightenments, we will have to save Hume and Pufendorf for any follow-up research.

the chapters of this thesis have been structured in order to allow them to be read as isolated interpretations of five early-modern philosophers.

At the same time, we may provide several reasons to justify our selection of authors. Firstly, as we have already explained at the start of this introduction, these five thinkers share a number of fundamental presuppositions. Consequently, they develop their moral theories from a similar point of departure. Furthermore, all five thinkers are at least to some extent concerned with epistemology. With the possible exception of Condillac, they also make important contributions to moral philosophy. This combined interest in epistemology and ethics makes these thinkers especially suited to our discussion of moral epistemology. Finally, despite their shared preconceptions, these five thinkers each have their own goals and preoccupations that will provide diversity to the scope of our discussion. For instance, Condillac's interest in the role of words in cognition will provide us with the opportunity to explore at length the relation between language and morality. Likewise, Helvétius's engagement in contemporary polemics will allow us to illustrate how a theory of moral epistemology could also provide the basis for an intervention in more concrete social and political debate.

As the chapters of this thesis have diverse preoccupations, we will also employ slightly different methodologies in our investigation. In most chapters, we will follow the methodological suggestions of Knud Haakonssen on the interpretation of early-modern moral philosophy. Haakonssen presents an alternative to the popular methodology championed by Quentin Skinner and the Cambridge School, which considers all historical utterances – including philosophical treatises – as performative interventions in contemporary debate. Accordingly, proponents of the Cambridge school primarily seek to interpret their source material by placing it in the context of contemporary social and political discussions.[17] Haakonssen maintains that while this is a viable approach for many kinds of source material, our understanding of early-modern moral philosophy would also benefit from an investigation focused on philosophical ideas themselves.[18] Following Haakonssen suggestions, we will therefore explain

[17] Quentin Skinner, 'Meaning and Understanding in the History of Ideas', *History and Theory* 8 (1969) 3-53 contains the classic formulation of this methodology.

[18] Haakonssen, *Natural Law and Moral Philosophy* 10: 'it would seem that to be part of the intellectual historian's task to write the history of the utterance not only as a performance but also as a reference. The latter, however, cannot be done except through an investigation of the

the positions of Hobbes, Locke, Condillac and Rousseau primarily by providing a detailed reconstruction of the arguments of their moral philosophy. Yet our goal is not always to provide a fully coherent reconstruction of the philosophy of Hobbes, Locke, Condillac, and Rousseau. Rather, our interpretation will be perceptive to any fissures and inconsistencies in the arguments of these thinkers. We will therefore proceed upon the assumption that uncovering and displaying the loose ends of a theory will ultimately be more instructive than trying to resolve all inconsistencies to produce a fully coherent interpretation. Finally, as it is clear that the purpose of Helvetius' writings was not merely the discussion of abstract philosophical theory, but also intervention in contemporary polemics, our chapter on Helvétius will occasionally supplement this focus on philosophical arguments with a more contextual approach to interpretation.

Contribution to Scholarship

Throughout our investigation, we will engage with scholarly literature on several levels. As this thesis features largely self-sufficient interpretations of Locke, Hobbes, Condillac, Helvétius, and Rousseau, we will predominantly engage with scholarly literature dedicated to the interpretation of their individual philosophies. The chapters on Hobbes, Locke, and Rousseau, will therefore feature detailed discussions of the extensive range of scholarly interpretations of their writings. Our discussion of their moral epistemology will then allow us to evaluate some of the main lines of interpretation of their moral philosophy as a whole. Yet as we have seen, this thesis will also feature some recurring issues and questions. Accordingly, we will also engage with the literature on some more general themes. In particular, we will contribute to our understanding of two distinct issues that have been either underrepresented or misconstrued in the literature on the interpretation of early-modern moral philosophy, namely regarding the origin of normativity and the role of the passions in practical deliberation and morality.

To begin with, our combined discussion of epistemology and ethics will allow us to reconsider the source of normativity in early-modern moral philosophy. More specifically, we will be able to reconstruct the relation between facts and values in early-modern moral philosophy. Throughout our main

purported objects of reference, which, in intellectual history, will primarily be the ideas employed by an historical speaker in making an utterance.' Paulien Westerman, *The Disintegration of Natural Law Theory* 6-11 endorses a similar methodology for the history of moral philosophy.

discussion, we will find that some scholars have simply assumed that early-modern thinkers accept Hume's famous prohibition against deriving values from facts.[19] Yet our own interpretation will suggest that Hobbes for instance did not yet recognise the *is/ought* distinction. In fact, as Hobbes infers natural law from a conception of human nature and generalisations about the human condition, we will maintain that his aim was precisely to derive morality from certain distinct properties of empirical reality.[20] Subsequently, we will see that the distinction between facts and values does arise in Locke's philosophy as a consequence of his dichotomy between ideas of empirical objects acquired by means of experience and moral concepts constructed by the mind at will. At the same time, we will discover that strict adherence to this epistemological dichotomy obscures the source of normativity in Locke's account of moral reasoning. Our discussion of moral philosophy in the context of epistemology will then suggest that the *is/ought* dichotomy is closely related to an empiricist account of human understanding.[21]

Furthermore, our investigation will provide a different perspective on the relation between reason and the passions in early-modern moral philosophy. As we will see, several modern interpreters have ascribed to Hobbes, Locke, and

---

[19] Stephen Darwall, 'Normativity and Projection in Hobbes's Leviathan', *The Philosophical Review* 109 (2000) p.313-347 for instance assumes that Hobbes ascribed to the distinction.

[20] Hume's famous prohibition against deriving values from facts has been interpreted by many commentators, including J.L. Mackie, *Hume's Moral Theory* (London 1980) 47, as directed at moral realists like Cudworth and Clarke. Yet Sophie Botros, *Hume, Reason and Morality: A legacy of contradiction* (London and New York 2006) 72-95 shows that Hume's is/ought distinction is not only directed at moral realism, but also at moral naturalists like Hobbes, who would aim to derive normative conclusions from empirical facts.

[21] Westerman, *The Disintegration of Natural Law Theory* 288-289 also relates the emergence of the fact/value dichotomy to an empiricist theory of knowledge. Yet this relation is also suggested by the fact that the dichotomy is rejected by modern detractors of empiricism. Thus pragmatists like Hilary Putnam reject the notion that we may derive a theory of knowledge from a descriptive account of human understanding. Accordingly, in his *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge MA 2002), Putnam tries to demonstrate that the *is/ought* dichotomy rests on an unfounded distinction between ideas of empirical reality and moral concepts. Earlier, in his *Reason, Truth and History* (Cambridge 1981) 127-149 Putnam had shown that the dichotomy between facts and values cannot be sustained, as scientific practice already relies on a number of values, the most basic of which is that the aim of science is to establish truth. Alternatively, conservative thinkers like Alisdair MacIntyre, *After Virtue: A Study in Moral Theory* (London 2011 [1981]) 69-71 present the dichotomy as a consequence of the empiricist rejection of Aristotelian teleology. As MacIntyre seeks to revive a teleological conception of morality, he also rejects the fact/value distinction.

Rousseau the view that the purpose of reason in practical deliberation is solely to find the means towards the satisfaction of our desires. Accordingly, these scholars maintain that Hobbes and Locke anticipated the Humean view that 'reason is, and ought to be, the slave of the passions.'[22] Yet our own investigation of the account of practical deliberation of these thinkers will demonstrate that by weighing the consequences of our actions, reason not merely finds the means to whatever desires we might experience, but may also participate in determining the ends of our pursuits. While our reading assigns reason a more prominent role in practical deliberation, it also explains how reason may determine the morality of our actions on the basis of their consequences. Accordingly, we will interpret the moral philosophies of Hobbes, Locke, Condillac and Helvétius as early examples of consequentialism, thereby substantiating the suggestion that the moral philosophies of these thinkers anticipate the preconceptions of classical utilitarianism.[23]

[22] Hume, *A Treatise of Human Nature* ed. by David Fate Norton and Mary J. Norton (Oxford 2000) II.iii.3 §4 p.266.

[23] The role of Hobbes, Locke, and in particular Helvétius as forerunners to classical utilitarianism is attested by: Colin Heydt, 'Utilitarianism before Bentham', in: Ben Eggleston and Dale E. Miller (eds.), *The Cambridge Companion to Utilitarianism* (Cambridge 2014) 16-37; Frederick Rosen, *Classical Utilitarianism from Hume to Mill* (London and New York 2003) 82-96; Pierre Force, *Self-Interest before Adam Smith, A Genealogy of Economic Science* (Cambridge 2003) 93-94; Albert Hirschmann, *The Passions and the Interests: Political Arguments for Capitalism before its Triumph* (Princeton 1977).

1. Reasoning about Consequences: Hobbes's *Consequentialist* Theory of Morality

Introduction

The moral philosophy of Thomas Hobbes is based on a comprehensive investigation of human nature. The most paradigmatic formulation of Hobbes's science of human nature can be found in the first part of *Leviathan*.[24] Like any other science concerned with physical objects, the method of this science of human nature is empirical.[25] Although building upon the findings of more fundamental scientific fields of inquiry, Hobbes thinks our knowledge of human nature derives both from observation of human behaviour anfd introspection of our own mental states.[26] On the basis of the findings of his own investigation of human nature, Hobbes first describes the cognitive processes by which humans are able to reason about empirical objects and their consequences on the basis of sense experience and with the help of language. This description also has normative implications, in the sense that it indicates which forms of reasoning are valid, and which methods lead us into error or absurdity. Secondly, Hobbes provides an account of human motivation, indicating that all our actions proceed

[24] In this chapter, we will generally follow the suggestion by William Sacksteder, 'Hobbes's Science of Human Nature', *Hobbes Studies* 3 (1990) 35-53, 37 that part I of *Leviathan* contains the definitive synthesis of Hobbes's science of human nature. To be sure, Hobbes discusses many specific aspects of human nature more extensively in some of his other writings – specifically in *De corpore, De homine* and *Questions Concerning Liberty, Necessity and Chance.* In the following, we will therefore occasionally turn to these latter works, but only for the purpose of explaining and contextualising the often concise arguments we find in the *Leviathan*.

[25] In contrast to some antiquated readings, including F.S. McNeilly, *The Anatomy of the Leviathan* (New York 1968) 66-87 and Aloysius Martinich, *Thomas Hobbes* (London 1997) 98 who claim that Hobbesian science is one long deduction, with the conclusions of more applied subjects like morals and politics derived entirely from the results from more fundamental fields like physics and geometry. This notion is definitively refuted by William Sacksteder, 'Three Diverse Sciences in Hobbes: First Philosophy, Geometry, and Physics', *Review of Metaphysics* 45 (1992) 739-772 and Tom Sorell, *Hobbes* (London and New York 1986) 5-6, who maintain that Hobbes favours an empiricist methodology in the natural sciences, which encompass the science of human nature.

[26] Hobbes, *Leviathan* ed. by Edwin Curley (Indianapolis and Cambridge 1994) Introduction §3 p.4: 'whosoever looketh into himself and considereth what he doth, when he does *think, opine, reason, hope, fear,* &c, and upon what grounds, he shall thereby read and know, what are the thoughts and passions of all other men upon the like occasions.' See also XXXII.1 p.245; Sacksteder, 'Hobbes's Science of Human Nature'; Laurens van Apeldoorn, 'Hobbes on the Scientific Study of the Human Mind', *Archiv für Geschichte der Philosophie* 97 (2015) 308-333; Sorell, *Hobbes* 7; 25-27.

from certain passions and desires. Yet while the passions ultimately provide motivation, it is the role of reason to decide which desires are worth pursuing, and to find the means to their fulfilment.

After outlining man's capacities and motivations, Hobbes devises a theory of morality that follows from this conception of human nature. This theory of morality is presented in the form of natural laws – basic moral precepts that are both recognisable and applicable to all human beings. As the first step in his deduction of the laws of nature, Hobbes conducts his famous thought experiment of the state of nature, posing the question how human beings would behave towards each other without any established authority to keep them in check. Hobbes describes how in such a state, the inherent and unrelenting desire for pleasure combined with the capacities for speech and reasoning would bring human beings to seek power, riches, and glory – even to the detriment of others. As a result, competing individuals would inevitably come into conflict with each other. Hobbes then famously concludes that in view of the abilities and inclinations of human nature, life in such a state would be 'solitary, poor, nasty, brutish, and short.'[27] Yet while the war of all against all that characterises the state of nature is a consequence of the combination of man's sagacity with his destructive tendencies, Hobbes maintains that human nature also contains the intellectual abilities to overcome this adverse situation. Thus Hobbes states that: 'reason suggesteth convenient articles of peace, upon which men may be drawn to agreement. These articles are they which otherwise are called the Laws of Nature.'[28]

Hobbes thus claims that natural law – and therefore a basic yet universal formulation of morality – can be deduced by employing our capacity for reasoning.[29] Yet upon Hobbes's own account of reason, it is not immediately evident in what manner reason would arrive at the normative guidelines of natural law. Hobbes maintains that reason is an artificial faculty dependent on language, thereby rejecting the claim of the Scholastics and arguably even Hobbes's immediate predecessor Grotius that human beings simply have an intuitive or *a priori* access to the verities of natural law.[30] Instead, Hobbes

[27] Hobbes, *Leviathan* XIII.9 p.76.

[28] Hobbes, *Leviathan* XIII.14 p.78.

[29] Hobbes, *Leviathan* XV.40 p.100: 'the true doctrine of the laws of nature is the true philosophy.'

[30] The interpretation of Grotius advanced here is contested by Richard Tuck, 'The 'modern' theory of natural law', 99-120 who portrays Grotius as prefiguring Hobbes's concept of natural law

maintains that reasoning is nothing but drawing conclusions from combinations of words that serve as marks to ideas of empirical objects.[31] But how can such a form of reasoning by itself lead to any normative conclusions? Hobbes of course maintains that natural law is a rule indicating the most effective means towards self-preservation – thereby indicating that self-preservation should be the prevailing goal of our actions. But Hobbes himself does not spell out the deduction by which he reaches this normative conclusion. This raises the question why self-preservation should be the purpose of natural law. Why not felicity, conceived as the continuous satisfaction of whatever desires individuals might have?[32] This question acquires even greater relevance in light of the common critique that Hobbes's conception of natural law based solely on the desire for self-preservation leads to a society that provides peace and safety, but does very little to promote the happiness of its citizens.[33] In light of these issues, we will investigate Hobbes's philosophy starting from the question: how may reasoning with words attached to ideas acquired by means of the senses arrive at the normative conclusion that natural laws are guidelines to our self-preservation?

Modern interpreters of Hobbes have formulated a variety of answers to these questions. A first line of interpretation maintains that natural laws are in fact not the product of reasoning upon words attached to ideas, but rather follow from a 'natural reason' innate to human nature. According to scholars like Gert and Hoekstra, this 'natural reason' is then able to identify which actions and pursuits are inherently rational, of which the concern for self-preservation is the

---

solely based on self-preservation. Yet as Robert Shaver, 'Grotius on Scepticism and Self-Interest', *Archiv für Geschichte der Philosophie* 78 (1996) 27-47 and Westerman, *The Disintegration of Natural Law* 132-140 have pointed out, Tuck's interpretation is based mainly on Grotius's *De iure praedae*, and early work not published until the nineteenth century. In *De iure belli ac pacis*, Grotius' main work, he departs from his earlier views towards a more Aristotelian position.

[31] Whether this is also applies to mathematics and metaphysics is an issue disputed among commentators, but not particularly relevant for our purposes. Cf. Sorell, *Hobbes* 55-67.

[32] Hobbes, *Leviathan* XI.1 p.57 defines felicity as: 'a continual progress of the desire, from one object to another, the attaining of the former being still but the way to the latter.'

[33] Alisdair MacIntyre, *A Short History of Ethics* (Oxford 2002 [1967]) 133-134. Rousseau in his *On the Social Contract* in: idem, *On the Social Contract, with Geneva Manuscript and Political Economy* ed. by Roger Masters, trans. by Judith Masters (Boston 1978) I.iv p.49 voices a variation on this critique, stating: 'It will be said that the despot guarantees civil tranquillity to his subjects. [...] Life is tranquil in jail cells too. Is that reason enough to like them?'

prime example.[34] This reading thus attributes Hobbes a conception of reason not unlike the innate faculty assumed by the scholastics and arguably Grotius as well. An alternative though related view has been advanced by Stephen Darwall, who ascribes to Hobbes a projectivist theory of morality. Darwall does not maintain that the normative priority of self-preservation follows from natural reason, but rather that human beings are naturally disposed to reason from the premise that self-preservation should be the purpose of their actions. This disposition then causes humans to project normative properties upon the objects they encounter through sensation.[35] A third line of interpretation, proposed by John Deigh, takes a position diametrically opposed to the view that natural law as guidelines for self-preservation is somehow innate to human understanding. Deigh instead maintains that when Hobbes claims that reasoning is '*reckoning* (that is, adding and subtracting) of the consequences of general names agreed upon,' he literally means that reasoning is determined by the conventional signification of words.[36] According to Deigh, the laws of nature ensuring self-preservation therefore follow from the composition of our language.[37]

However, these three interpretations dissent from what is by far the most common reading among scholars. The standard interpretation starts out from the premise that for Hobbes all human actions are motivated by the passions, and that in practical deliberation reason has the role of finding the *means* towards the actions desired. Because being alive is a necessary prerequisite to the satisfaction of any desires, the pursuit of self-preservation is also a conclusion of reason finding the *means* towards any of the objects of our passions. According to this interpretation, the natural laws designed to safeguard our self-preservation are the product of reason as means-to-ends thinking. In line with Hobbes denial that there is such a thing as a 'rational appetite,' reason would not participate in identifying which desires to pursue, but merely indicate

[34] Bernard Gert, 'Hobbes on Reason', *Pacific Philosophical Quarterly* 82 (2001) 243-257; Kinch Hoekstra, 'Hobbes on Law, Nature, and Reason', *Journal of the History of Philosophy* 41 (2003) 111-120, 118-120. Richard Tuck, 'Hobbes's Moral Philosophy', in: Tom Sorell (ed.), *The Cambridge Companion to Hobbes* 175-207, 187 approaches the same position by claiming that this focus on self-preservation is simply self-evident – an interpretation that does not explain much.

[35] Darwall, 'Normativity and Projection in Hobbes's Leviathan.'

[36] Hobbes, *Leviathan* I.v.2 p.22-23.

[37] John Deigh, 'Reason and Ethics in Hobbes's *Leviathan*', *Journal for the History of Philosophy* 34 (1996) 33-60. For a critical response to Deigh's interpretation see: Mark Murphy, 'Desire and Ethics in Hobbes's *Leviathan*: A Response to Professor Deigh', *Journal for the History of Philosophy* 38 (2000) 259-268.

the most effective method to achieve them.[38] Thus when Hobbes states that the laws of nature are 'dictates of reason,' he maintains that they are rational guideline outlining the surest *means* to felicity.[39]

Throughout this chapter, we will engage with these various interpretations more extensively. We will then see that none of them fully represents Hobbes's theory. Our own interpretation of Hobbes's conception of normative reason fundamental to natural law will expand upon the recent work of van Apeldoorn and Abizadeh.[40] This interpretation is based on a reading of Hobbes's account of practical deliberation that recognises that we may use reason to determine whether our passions are truly desirable. This reading thereby contests the common view that Hobbes regards the passions as given and that the role of reason in practical deliberation is limited to finding the means to whatever we desire.[41] Our own interpretation assigns reason a more active role in practical deliberation, as it also participates in the scrutiny and selection of our desires by considering the consequences of objects or actions. We will see that this reading is fully consistent both with Hobbes's claim that reason is an acquired rather than a natural faculty, and with his position that all pursuits following from the will ultimately derive from the passions. On the basis of this reading of Hobbes's account of reason and its role in practical deliberation, we will then be able to formulate a more accurate answer to the question how Hobbes thinks reason is able to arrive at normative conclusions – including those of natural law. Finally, we will outline how reason deliberating on the consequences of the passions inherent in human nature concludes that the desire for self-preservation has normative priority over all other desires.

---

[38] Hobbes, *Leviathan* I.vi.53 p.33. Cf. *De cive* I.iv.1 ann.

[39] Noel Malcolm, 'Hobbes and Spinoza', in: idem, *Aspects of Hobbes* (Oxford 2002) 27-52, 32. Jeffrey Barnouw, 'Reason as Reckoning: Hobbes's Natural Law as Right Reason', *Hobbes Studies* 21 (2008) 38-62; David Gauthier, *The Logic of the Leviathan: The Moral and Political Theory of Thomas Hobbes* (Oxford 1969) 21; Stephen Darwall actually endorsed the same interpretation in a book preceding the article discussed above: *The British Moralists and the Internal 'Ought'* 58-59.

[40] Arash Abizadeh, 'Hobbes on Mind: Practical Deliberation, Reasoning, and Language', *Journal of the History of Philosophy* 55 (2017) p.1-34; Laurens van Apeldoorn, 'Reconsidering Hobbes's Account of Practical Deliberation', *Hobbes Studies* 25 (2012) 143-165. Adrian Blau, 'Reason, Deliberation, and the Passions', in: A.P. Martinich and Kinch Hoekstra (eds.), *The Oxford Handbook of Hobbes* (Oxford 2016) 195-217 provides a similar reading.

[41] MacIntyre, *A Short History of Ethics* 134; Michael Allen Gillespie, *The Theological Origins of Modernity* (Chicago 2008) 236.

We will start our inquiry with a review of Hobbes's account of basic human cognition, moving subsequently to a discussion of his theory of reason as dependent on language. Meanwhile, we will also scrutinise John Deigh's interpretation that ascribes to Hobbes the view that the precepts of natural law are contained within the conventions of language. Subsequently, we will analyse and dismiss the reading that Hobbes thinks that natural law follows from an innate faculty of normative reason distinct from descriptive reasoning based on speech. After considering the interpretation of natural law as the means to our ends, we will then develop our own reading by first reconsidering Hobbes's account of practical deliberation, as well as his theory of subjective normative judgment, concluding that the latter is ultimately consequentialist in nature. Subsequently, we will examine to what extent the normative judgment of individuals is capable of rendering them sociable creatures, and conclude that it are ultimately the limits of subjective reasoning that necessitate the erection of a sovereign to make humans comply with natural law. Finally, we will conclude that Hobbes's deduction of natural law employs a method similar to that of practical deliberation. Thus by scrutinising the desires inherent in human nature in view of their consequences, Hobbes concludes that the desire for self-preservation takes priority over all other passions, and should consequently be the basis of natural law.

Experience and Understanding

The fundamental thesis of Hobbes's theory of human understanding, appearing right in the first chapter of the *Leviathan,* is that the content of all thoughts may ultimately be traced back to sense perception.[42] Hobbes describes how representations of objects arise in the mind due to pressure exerted upon the organs of sense. While sensation itself is fleeting, perceptions leave imprints upon the mind that give rise to what Hobbes calls the *imagination*. Describing the imagination as nothing but 'decaying sense,' Hobbes details how it preserves our sensations, thereby enabling what is commonly called memory. Over time, the

[42] Hobbes, *Leviathan* I.2 p.6; *De corpore* ed. by William Molesworth (London 1839) I.vi.1 p.66. Sorell, *Hobbes* 86-87 suggests that this empiricist position appears to conflict with some of Hobbes's more detailed views on the origin of knowledge, which indicate that in fact Hobbes thinks that many parts of the sciences – such as those concerning magnitude and motion – are not conducted through empirical enquiry. Cf. Marcus Adams, 'Hobbes, Definitions and Simplest Conceptions', *Hobbes Studies* 27 (2014) 35-60 for a contrary reading.

sensations in our imagination will decay and become more obscure because they are supplemented and replaced by the steady stream of new perceptions flowing from the senses. The resulting collection of sense-perceptions retained by the imagination is then what Hobbes calls *experience*.[43]

On the basis of this experience alone, it is already possible to make very basic kinds of judgments about the world around us. First of all, the imagination can perceive the differences and similarities between objects of sense and memory. This comparison of ideas does not require any independent and assenting faculty, as Hobbes thinks that judgment is nothing but *sensing* the similarities and differences between ideas and sensations.[44] Secondly, Hobbes explains that our perceptions are always acquired in an order that mirrors the sequence of phenomena perceived. Our memories therefore exhibit conformity to the temporal order of nature, as 'those motions that immediately succeeded one another in the sense continue also together after sense.'[45] This succession of memories reflecting the order nature then allows the mind to conjecture causal relations between objects and events.[46] Hobbes maintains that this form of understanding of the conjunctions between one object and another is common to both man and beast. Consequently, all creatures endowed with the abilities for sense-perception and imagination have the ability to understand causes and foresee potential effects, provided they have past experience of objects similar to the one they encounter. This form of conjecture is what Hobbes calls *prudence*, and can be used by both man and animal to guide their actions and satisfy their desires.[47] The extent of our prudence – and thus our ability to achieve our aims – then depends on the range of our experience, 'the *future* being but a fiction of the

[43] Hobbes, *Leviathan* II.2-4 p.8-9; *De corpore* IV.xxv.7 p.396.

[44] Hobbes, *De corpore* VI.xxv.8 p.399. Both Susan James, *Passion and Action: the Emotions in Seventeenth Century Philosophy* (Oxford 1999) 128 and Sorell, *Hobbes* 85 note that Hobbes sometimes ascribes the imagination cognitive abilities that go beyond merely sensing, recalling and distinguishing particular ideas. In other words, it is not clear that according to Hobbes's own description, the imagination would on its own have the capacity for some of the forms of thinking Hobbes ascribes to it.

[45] Hobbes, *Leviathan* III.2 p.12.

[46] As Sorell, *Hobbes* 30 n.1 notes, this position may appear to anticipate Hume's sceptical theory of causality. Yet this only holds for the prudential understanding of causality based solely on experience. Hobbes also thinks that humans do have the ability to ascertain causal relations with certainty through ratiocination – a theory dismissed by Hume as incomprehensible. Cf. David Hume, *An Enquiry concerning Human Understanding* ed. by Peter Nidditch (Oxford 1975) IV p.25-39.

[47] Hobbes, *Leviathan* III.7 p.13-14.

mind, applying the sequels of actions past to the actions that are present; which with most certainty is done by him that has most experience.'[48]

Hobbes therefore claims that the combination of sense and imagination alone already facilitates a rudimentary form of cognition that is not only behind all animal behaviour, but accounts for many human actions as well. Yet Hobbes also indicates that an understanding of the world merely based on prudence is both severely limited and prone to be inaccurate. He provides several reasons why this is the case. First of all, prudence is based solely on our own experience, excluding the possibility of benefitting from the wisdom of others. Secondly, the 'decaying sense' of our imagination on which prudence depends is fleeting, as it is continually supplemented and replaced by new experiences. Thirdly, Hobbes seems to think that prudence alone cannot ascertain causal relations with any certainty, as the relations perceived between objects of experience may not always be the same.[49] Finally, Hobbes claims that the imagination can only acquire and consider particular ideas based on singular experiences. By itself, it is unable to construct universal or abstract concepts that represent multiple objects at the same time. On the basis of prudence, humans and animals can therefore only deliberate and reflect on the singular objects and events they have experienced in the past.[50]

Language and Reason

While prudence is common to both man and beast, Hobbes claims that the higher cognitive ability of humans is due entirely to our capacity for language.[51] In Hobbes view, language thus not only has a communicative function that allows us to learn from others, it also changes the way we think, as we 'turn the reckoning of the consequences of things imagined in the mind into a reckoning of

[48] Hobbes, *Leviathan* III.7 p.14.

[49] Hobbes, *Leviathan* III.2 p.12; XLVI.2 p.454. It is not entirely clear how Hobbes thinks that a diversity of experiences would in fact confuse our prudential comprehension of causal relations. Possibly his point is that without language as a tool to accurately register the causal relations perceived, the mind would have difficulty distinguishing the merely casual from constant conjunctions.

[50] Hobbes, *Leviathan* IV.6 p.17; V.5 p.23-24.

[51] Michael Losonsky, *Linguistic Turns in Modern Philosophy* (Cambridge 2006) 49-50 completely disregards this aspect of Hobbes's philosophy of language. Losonsky claims that Hobbes's discussion of language is solely meant as a methodological prelude to his main project of political philosophy. Hopefully, our own discussion will demonstrate that Losonsky underestimates the importance of Hobbes's philosophy of language.

the consequences of appellations.'[52] According to Hobbes, it is therefore only by acquiring speech that we transcend mere prudence and develop our capacity for reasoning. Hobbes thinks that the imposition of words extends the fundamental abilities of the mind in a number of ways. First of all, words serve as marks to thoughts, thereby enhancing our otherwise fleeting memory of causal relations. Thus Hobbes thinks that words and the connections between them allow the 'registering of the consequences of our thoughts, which being apt to slip out of our memory and put us to a new labour.'[53] By registering our experience of causal relations, we then also increase our ability to distinguish and compare them, giving rise to more accurate judgment.

More importantly still, language also broadens the scope of human understanding by enabling us to think about species of things, rather than solely about individual objects. Hobbes is a nominalist regarding universal concepts, which means that he thinks that universals are not mental concepts, but rather terms with a certain definition:

> Of names, some are *proper*, and singular to one only thing, as *Peter, John, this man, this tree;* and some are *common* to many things, as *man, horse, tree*, every of which, though but one name, is nevertheless the name of divers particular things, in respect of all which together it is called an *universal*, there being nothing in the world universal but names; for the things named are every one of them individual and singular.[54]

As we have seen, Hobbes thinks that by itself, the imagination can only consider singular objects. But by using universal terms, human beings gain the ability to reason and communicate about whole classes of objects. The reference of these universal terms is then based on their definition.[55] With the help of these universal terms, humans can greatly expand and streamline their thinking. Hobbes himself gives the example of a person studying a triangle and discovering

---

[52] Hobbes, *Leviathan* IV.9 p.18.

[53] Hobbes, *Leviathan* IV.3 p.16. Cf. *De corpore* I.vi.11 p.79-80.

[54] Hobbes, *Leviathan* IV.6 p.17. Cf. *De corpore* I.ii.9 p.20; I.v.8 p.60: 'they err, that say the *idea of anything is universal*; as if there could be in the mind an image of a man , which were not the image of some one man, but a man simply, which is impossible; for every idea is one, and of one thing.' For a critical discussion indicating some of the implications and problems of this view, see G.K. Callaghan, 'Nominalism, Abstraction, and Generality in Hobbes', *History of Philosophy Quarterly* 18 (2001) p.37-55.

[55] Hobbes, *Leviathan* IV.12 p.19.

one of Euclid's theorems. Without the use of words, this person would have to repeat the mathematics every time she encounters a slightly different three-cornered object. But after acquiring and learning the definition of the universal term 'triangle,' this person would become able to attribute Euclid's theorem to any three-cornered object she encounters, without having to repeat her reasoning. Hobbes therefore thinks that with the help of universal terms, 'the consequence found in one particular comes to be registered and remembered as a universal rule [...] and delivers us from all labour of the mind, saving the first, and makes that which was found true *here* and *now*, to be true in *all times* and *places*.'[56] Universal terms then allow us to formulate general rules and propositions indispensible to reasoning and the practice of science.[57] Without universals, all our reasoning and knowledge would be of particulars.[58] It is therefore only by using well-defined universal terms that we can formulate laws of nature of both the moral and scientific kind.

According to Hobbes, universal terms, as well as proper names – together with a range of other types of words, including abstract terms, numerals, negations, and the copula – therefore transforms and enhance human understanding, thereby producing reason.[59] As these terms are related to each

[56] Hobbes, *Leviathan* IV.9 p.18.

[57] It is important to note that for Hobbes, the term 'science' does not exclusively refer to the systematic investigation of empirical reality by means of the scientific method. This narrow usage of the term 'science' became only commonplace during the nineteenth century. Instead, for Hobbes, the term 'science' refers to any form of knowledge in which we employ universal terms, regardless of whether this knowledge has been acquired by means of what later has become known as the scientific method.

[58] Philip Pettit, *Made with Words: Hobbes on Language, Mind, and Politics* (Princeton University Press: Princeton 2008) 42-43 agrees that reason becomes possible only with the use of language. This is disputed by Abizadeh, 'Hobbes on Mind', 18, who points out that Hobbes, *Leviathan* V.5 p.23 himself envision the possibility that 'man reckons without the use of words, which may be done in particular things.' Yet as we have seen, this reckoning regarding particular things is more properly described as a form of prudence rather than reason, which by its very nature reckons with universal terms leading to general conclusions.

[59] Hobbes, *Leviathan* IV.16 p.20; *De corpore* I.iii.3-4 p.32-34. Besides universal terms, Hobbes identifies a number of additional types of terminology that expand the scope of our reasoning. While it is beyond our present purposes to provide a catalogue of all these types of words, we may briefly note the function of abstract terms. Hobbes describes how this type of names allows us to communicate and reason about specific aspects or qualities of objects, independent from the idea of the object itself. Hobbes's examples include terms like 'motion,' 'heat,' and 'length.' These do not refer to any specific object in nature, and neither do they signify an idea in the mind of a speaker. Instead, an abstract term refers to a property common to many different objects, while its signification derives from a definition constituted or accepted by speakers themselves. Like

other by grammatical and logical relations, they collectively provide a framework for reasoning.[60] Thus while animal or prudential thinking is entirely determined by experiences assessed through innate cognitive structures, humans order their thoughts through the artificial structure of language. Consequently, human reasoning is both broader in scope and detached – although not entirely dissolved – from the particular order of ideas acquired through experience. Yet Hobbes is not altogether clear on the precise implications of this transformation of human understanding by means of language. Consequently, scholars have debated whether Hobbes'ss account of the role of language in cognition is sufficient to explain that human beings are conscious, reflective, and autonomous agents rather than merely passive beings conditioned by their environment.[61] Yet as this issue is not of primary importance to our own discussion, we may refrain taking a position in this debate.

Language and Natural Law

A more relevant issue for our purposes is the question to what extent language not only constitutes reasoning, but also determines its outcome. In the preceding, we have seen that reasoning depends on both universal and abstract terms. Yet we have also seen that Hobbes thinks all our ideas are particular. Universal and abstract terms therefore do not signify any universal or abstract conceptions in the mind – simply because Hobbes thinks these cannot exist. The meaning of universal and abstract terms is rather determined by their definition, which indicates the extent of their reference to objects of experience. It therefore follows that when we reason with general and abstract terms, the outcome of our reasoning will be determined by the definitions of the terms we employ. This is also the reason why Hobbes emphasises repeatedly that our reasoning should

---

universal terms, abstract names enable forms of thought not restricted to the consideration of ideas of individual objects acquired through sensation. Therefore they allow humans to communicate and reason about objects of experience in a way that transcends merely prudential thinking.

[60] Hobbes, *Leviathan* II.10 p.11. Cf. Sorell, *Hobbes* 30-32 on this notion of language and reason not as the replacement of prudence, but as a way of structuring and connecting knowledge gained through experience.

[61] Abizadeh, 'Hobbes on Mind', 19-30 does think that the extent of language is such that humans escape passivity and gain the ability to assent consciously to statements and propositions. Cf. Michael Losonsky, *Enlightenment and Action from Descartes to Kant* (Cambridge 2001) 42-70.

commence with scrutinising and settling the definitions of the terms we aim to use.[62]

Yet this importance of definitions to our reasoning also raises the question from where Hobbes thinks these definitions of universal and abstract terms are derived. On first inspection, his answer appears to suggest that the definitions fundamental to our reasoning are nothing but human conventions:

> From hence also this may be deduced, that the first truths were arbitrarily made by those that first of all imposed names upon things, or received them from the imposition of others. For it is true (for example) that *man is a living creature*, but it is for this reason, that it pleased men to impose both those names on the same thing.[63]

Hobbes's writings abound with similar passages in which he seems to endorse the view that the meaning of words depends on arbitrary conventions determined by the inventors of language. Elsewhere, he states that 'names have their constitution, not from the species of things, but from the will and consent of men.'[64] Consequently, some of Hobbes's readers – Leibniz among them – have claimed that Hobbes holds that man's reasoning is entirely determined by the conventional meaning of words. Hobbes would therefore retain a conventionalist theory of truth, according to which the truth and falsity of a propositions is entirely dependent on linguistic convention.[65]

This line of interpretation has recently been revived by John Deigh, who applies this reading to explain how Hobbes thinks that reason would deduce natural law. According to Deigh, moral philosophy producing natural law is a form of science that relies on reasoning with aptly defined universal terms. Furthermore, Deigh assumes that the proper definitions of words can be found among the linguistic conventions that make communication possible. Deigh then concludes that for Hobbes the outcome of both normative and scientific

---

[62] Hobbes, *Leviathan* V.20 p.26: 'the light of human minds is perspicuous words, but by exact definitions first snuffed, and purged from ambiguity; *reason* is the *pace*; increase of *science*, the *way*; and the benefit of mankind, the *end*.'

[63] Hobbes, *De corpore* I.iii.8 p.36; I.iii.9 p.37: 'Now *primary* propositions are nothing but definitions, or parts of definitions, and these only are the principles of demonstration, being truths constituted arbitrarily by the inventors of speech, and therefore not to be demonstrated.'

[64] Hobbes, *De corpore* I.5.1 p.56.

[65] McNeilly, *The Anatomy of the Leviathan* 66-76; Martinich, *Thomas Hobbes* 98; For a discussion of the passage from Leibniz cf. J.W.N. Watkins, *Hobbes's System of Ideas: A Study in the Political Significance of Philosophical Theories* (London 1965) 144.

reasoning is determined by the linguistic conventions upheld by the community of speakers. On this reading, the normative primacy of self-preservation over all other pursuits, as well as natural law in general, would simply follow from the constitution of our language. Deigh concedes that linguistic conventions are likely to reflect characteristics of relevant aspects of empirical reality – in this case presumably human nature. Nonetheless, the reasoning leading to natural law is based on artificial linguistic conventions drawn up by the founders of language and upheld by the community of speakers.[66]

The interpretation proposed by Deigh has not found many followers.[67] There are several reasons why Deigh's reading misrepresents Hobbes's views – two of which we will mention. First of all, Deigh's interpretation of natural law as following from linguistic conventions does not clarify why individuals would be compelled or obliged to follow its precepts. Given Hobbes'ss position that desires and appetites ultimately occasion all human action, natural laws would have to be formulated and presented as guidelines that facilitate the fulfilment of these desires in order for individuals to act on them.[68] In other words, natural law needs to be based on man's moral psychology to be in any way convincing and obliging to its recipients. Deigh's conventionalist interpretation rather presents natural law as a deontic rule deduced through reason, but then fails to explain how or why these natural laws would induce obedience if these laws are constructed on the basis of linguistic conventions rather than the interests of individuals.[69]

More importantly, Deigh's reading – as well as the conventionalist interpretation in general – relies on a misrepresentation of Hobbes's epistemology and philosophy of language. We have seen that Deigh ascribes to Hobbes the view that socially upheld linguistic conventions are the source of the definitions of the terms used in reasoning. Yet in Hobbes's own writings, we do not find anything like a theory of meaning detailing the role of linguistic conventions in circumscribing the definitions of terms.[70] In fact, Hobbes generally presents the process of imposing terms as a private process, in which

[66] Deigh, 'Reason and Ethics in Hobbes's *Leviathan*.'

[67] On the contrary, it is explicitly criticised by Murphy, 'Desire and Ethics in Hobbes's *Leviathan*.'

[68] Hence Hobbes, *Leviathan* XIII.14 p.78 states that natural laws follow from 'fear of death, desire of such things as are necessary to commodious living.'

[69] Murphy, 'Desire and Ethics in Hobbes's *Leviathan*', 265 makes the same point.

[70] Sorell, *Hobbes* 38 and Bernard Gert, 'Hobbes on language, metaphysics, and epistemology', *Hobbes Studies* 14 (2001) 40-58, 48 agree with this assessment.

individuals are free to devise their own definitions for the words they employ. Hobbes therefore thinks that in the first instance, the function of words is to serve as *marks* to our thoughts, while their use as *signs* in communication is secondary.[71] While Hobbes briefly mentions that speakers must adhere to some extent to conventions in order to communicate effectively, Hobbes never suggests that these linguistic conventions also influence or even determine the content and sequence of our private thoughts.

On the contrary, Hobbes thinks that individuals should *not* simply borrow the meaning of words from others, but rather 'examine the definitions of former authors, and either correct them, where they are negligently set down; or to make them himself.'[72] Individuals should therefore derive definitions 'from their own meditation.'[73] But what does Hobbes mean by this, and how can this statement be reconciled with his view that 'the first truths were arbitrarily made by those that first of all imposed names upon things?' In order to answer these questions, we must first recognise that for Hobbes, the truth of propositions is by no means entirely detached from our experience of empirical reality. Although Hobbes is decidedly vague on the issue, a number of scholars contesting the conventionalist reading have attempted to reconstruct how according to Hobbes, human beings derive their definitions of universal terms from experience. Thus James Watkins maintains that since Hobbes allows that common names can stand for an accident common to many things enables him to avoid a truly conventionalist theory of truth.[74] A more recent interpretation has been proposed by Marcus Adams, who argues that for Hobbes the science of geometry cannot be conventional, and thereby can provide a solid foundation for the development of the other sciences.[75]

For our present purposes, it is not necessary to discuss these readings in great detail or to decide which one is the more accurate. What matters is that

---

[71] Hobbes, *De corpore* I.vi.11 p.80: 'In the method of invention, the use of words consists in this, that they may serve for marks, by which, whatsoever we have found out may be recalled to memory [...] but not as signs by which we declare the same to others; so that a man may be a philosopher alone by himself, without any master.' Cf. Sorell, *Hobbes* 86.

[72] Hobbes, *Leviathan* I.iv.13 p.19.

[73] Hobbes, *Leviathan* I.iv.13 p.19.

[74] Watkins, *Hobbes's System of Ideas* 144-150.

[75] Marcus Adams, 'Hobbes, Definitions and Simplest Conceptions', *Hobbes Studies* 27 (2014) 35-60 maintains that the definitions used in geometry are derived from so-called 'simplest conceptions,' such as 'place' or 'motion' acquired through experience. Cf. David Gauthier, 'Hobbes on Demonstration and Construction', *Journal of the History of Philosophy* 35 (1997) 509-521.

these interpretations show that for Hobbes, the definition of terms – and thereby the truth of propositions – is not merely conventional, but ultimately depends on our experience of empirical reality.[76] But then how should we interpret Hobbes's conventionalist sounding claims that truth is the consequence of the arbitrary imposition of names? The most probable explanation is that Hobbes does not intend to claim that truth itself is but a collectively upheld fiction, but merely that the terms in which truth is expressed are arbitrary and conventional. His point is therefore that words themselves are nothing more than arbitrary signs whose relation to objects is entirely conventional. Hobbes thereby takes position against the view – first expressed in Plato's *Cratylus* but still common in the seventeenth century – that there is a natural relation between linguistic signs and the things they reference.[77]

In addition, Hobbes possibly also meant that the reference of universal terms is – at least to some extent – dependent on our own will. Thus we are free to choose the particular properties of experience by which we classify natural objects into species. But once defined, we are obliged to apply universals in line with our experience. Finally, when Hobbes states that truths are made by those who first imposed language, he may refer to the fact that truth is an exclusive property of linguistic propositions. Without language, mental discourse proceeds following innate cognitive structures. Only by rendering our thoughts in propositional form does it become possible to evaluate them as either true or false. In other words, Hobbes intended to say that language makes our thoughts truth-apt – a point he also makes elsewhere.[78] Following these explanations, it is almost certain that Hobbes did not hold a conventionalist theory of truth, and neither did he think that our reasoning producing natural law merely follows the conventions of language.

---

[76] Malcolm, 'Hobbes's Science of Politics and his Theory of Science', in: idem, *Aspects of Hobbes* 146-155, 152-153 and Sorell, *Hobbes* 41-42 agree with this assessment.

[77] Adams, 'Hobbes, Definitions and Simplest Conceptions', 53-55. Cf. Avi Lifschitz, 'The Arbitrariness of the Linguistic Sign: Variations on an Enlightenment Theme', *Journal of the History of Ideas* 73 (2012) p.537-557 for a broader discussion of anti-Cratylist arguments in the early-modern period.

[78] Hobbes, *Leviathan* IV.11 p.18-19: 'For *true* and *false* are attributes of speech, not of things. And where speech is not, there is neither *truth* nor *falsehood*. *Error* there may be, as when we expect that which shall not be, or suspect what has not been; but in neither case can a man be charged with untruth.' Cf. *De corpore* I.iii.7-8 p.35-36. For a more extensive discussion of the importance of truth-aptness in reasoning, cf. Abizadeh, 'Hobbes on Mind', 19-20.

Finally, this analysis is corroborated by considering Hobbes's more specific views on the ways in which language may influence our thoughts. Hobbes is in fact highly concerned with the possibility that opinions are swayed and reasoning perplexed by certain forms of language:

> Wherefore, as men owe all their true ratiocination to the right understanding of speech; so also they owe their errors to the misunderstanding of the same; and as all the ornaments of philosophy proceed only from man, so from man also is derived the ugly absurdity of false opinions. For speech has something in it like to a spider's web, [...] for by contexture of words tender and delicate wits are ensnared and stopped; but strong wits break easily through them.[79]

At first inspection, this passage again suggests that it is language itself that determines our reasoning and has the ability to lead our thoughts astray. If Hobbes really thought that conventional meanings used for communication shape our thinking, we would then expect him to be worried about inaccurate or imprecise linguistic conventions leading us into error. Yet if we inspect Hobbes's more detailed descriptions of the power of words to sway our thoughts, we find that he is mainly concerned about rhetorical, deceptive and absurd forms of speech. Thus Hobbes in particular warns against the ability of metaphorical language to trick us into believing false opinions.[80] In addition, he maintains that people can be deceived by the use of terminology without any signification. Examples include concepts like 'incorporeal substance' or 'free will,' as well as the inconceivable claim that there might be 'accidents of bread in cheese,' all of which cannot signify any idea in the mind and are literally inconceivable and absurd.[81] According to Hobbes, these vacuous terms are able to mislead us into accepting otherwise invalid argumentation – a practice he sees commonly employed by scholastic philosophers.[82] While Hobbes is therefore clearly

[79] Hobbes, *De corpore* I.iii.8 p.36.

[80] Hobbes, *Leviathan* I.iv.4 p.21. Hobbes subsequently discusses two more abuses of language not relevant for our discussion, namely lying and insulting.

[81] Hobbes, *De corpore* I.iii.1 p.30: 'there can be no succession of thoughts to answer them.' Cf. *Leviathan* I.iv.20-21 p.21

[82] Hobbes, *Leviathan* IV.21 p.21; XLVI.24-42 p.462-468; *De corpore* I.iii.4 p.34. Stewart Duncan, 'Hobbes, Signification, and Insignificant Names', *Hobbes Studies* 24 (2011) 158-178, 175-178 points out that most of Hobbes's examples of absurd or insignificant terms are only so within Hobbes's own materialist metaphysics. Duncan therefore states that Hobbes's critique of scholastic terminology ultimately boils down to a rejection of scholastic metaphysics and the concepts it employs.

sensitive to the fact that metaphorical and absurd speech may lead us astray, he never mentions the possibility that our thoughts may be distracted by the conventional meanings of words.

In the end, we must conclude that Deigh's interpretation of natural law as based on linguistic conventions does not stand up to scrutiny. Yet our effort in refuting it has also lead us to a general understanding of Hobbes's theory of reasoning as based on language. We have seen that Hobbes's own explication of this theory is not always easily understood. Yet for all its ambiguity, we have also seen that Hobbes's theory is most plausibly interpreted as a form of empiricism, as it describes how human beings use language to reason with universal and abstract terms ultimately defined though experience. Hobbes then thinks that by employing these well-defined terms in syllogisms, we would be able to produce the universal propositions of science. Yet while this interpretation of Hobbes's epistemology and philosophy of language avoids the conventionalism supposed by Deigh, it also raises new questions if we would consider it in light or our investigation into the basis of natural law in reason. Thus our interpretation of reason as grounded in experience explains the method for descriptive reasoning. But it does not yet incorporate an account of the normative reasoning required for deducing natural law.

Normative Reason as Innate?

Accordingly, several scholars have questioned whether reasoning conducted through language would in fact be able to arrive at the *normative* conclusions of natural law. These authors – Stephen Darwall and Bernard Gert among them – maintain that Hobbes's conception of reason as clarified above is exclusively concerned with descriptive propositions. Reason thus produces scientific laws that describe how the world functions, but by itself it cannot tell us what we should do or how we should live our life.[83] Thus reason may supply us with the means to our ends, but it cannot show us these ends themselves. These

[83] Darwall, 'Normativity and Projection in Hobbes's Leviathan', 313-319; Gert, 'Hobbes on Reason', 246-248: 'Those who take verbal reason, or even instrumental reason, to be the central concept in Hobbes's moral philosophy, should offer some account of how the benefit of mankind becomes the end or goal of science or reasoning, for it is quite clear that neither verbal reason nor instrumental reason can provide such an end on its own.' Cf. Hoekstra, 'Hobbes on Law, Nature, and Reason.'

interpreters thereby apply to Hobbes philosophy the dichotomy between descriptive and normative statements that has become famous as Hume's law.[84] They assume that Hobbes would not have violated the is/ought distinction by attempting to derive normative conclusions from empirical facts. Gert and Darwall therefore conclude that when Hobbes stated that natural law is a dictate of reason, he could not have meant that it is a dictate of reasoning with universal terms referencing objects of experience.

As an alternative, these authors propose that besides what they refer to as 'verbal reason,' Hobbes retains an additional conception of reason responsible for moral judgment. A cursory reading of Hobbes writings indeed yields some passages that seemingly suggest that human beings have an innate rational faculty that would allow them to intuitively arrive at the normative conclusions of natural law. Thus Hobbes states that '*reason*, which is the *law of nature* itself, has been given to each and every man directly by God as a Rule for his actions.'[85] Gert, supported by Kinch Hoekstra, therefore claims that natural law is a product of 'natural reason,' a faculty of moral judgment innate to human nature.[86] In a somewhat similar vein, Darwall maintains that according to Hobbes, 'we reason *from the end* of self-preservation itself, not from the fact that it is our end.'[87] Inspired by Hume's projectivist theories of morality and causality, Darwall attributes to Hobbes the theory that we project our innate inclination for self-preservation upon the objects of experience, similar to how we project our experience of colours. According to this interpretation, the projection of normativity upon the objects of experience is then the starting point for moral reasoning. While Gert and Darwall disagree upon the specific cognitive mechanisms at work, they both ascribe to Hobbes a theory in which morality and natural law follow from a certain faculty or property innate to human nature.

While Hobbes occasionally appears to suggest that the pursuit of self-preservation and adherence to natural law simply follow from an inclination or faculty innate to human nature, there is sufficient textual evidence suggesting the contrary. Thus Hobbes emphasises that 'reason is not, as sense and memory,

[84] Cf. Zagorin, *Hobbes and the Law of Nature* 112-117 for an alternative discussion of attempts to evaluate Hobbes's moral philosophy in light of Hume's is/ought distinction.

[85] Hobbes, *De cive* I.iv.1. Cf. *De corpore* I.i.1 p.1.

[86] Hoekstra, 'Hobbes on Law, Nature, and Reason' 118-120. Cf. Bernard Gert, *Hobbes, Prince of Peace* (Cambridge 2010) 48-53; Gert, 'Hobbes on Reason'.

[87] Darwall, 'Normativity and Projection in Hobbes's Leviathan' 318.

born with us, nor gotten by experience only, as prudence is, but attained by industry, first in apt imposing of names, and secondly by getting a good and orderly method.'[88] This passage clearly indicates that Hobbes considered reason an artificial faculty dependent on language. At the same time, Hobbes's statements that appear to refer to an innate faculty of normative reason independent from 'verbal reason' are often brief, and occur outside the context of Hobbes's dedicated discussions of reason. They may also be interpreted as merely indicating the normative function of artificial reason as dependent on language.[89] The textual evidence for the interpretation that Hobbes attributed to human beings an innate ability for normative reasoning separate from descriptive reasoning facilitated by language is therefore not compelling enough.

[88] Hobbes, *Leviathan* V.17 p.25. Cf. Hobbes, *Leviathan* I.viii.13 p.40: 'As for *acquired wit* (I mean acquired by method and instruction), there is none but reason, which is grounded on the right use of speech, and produceth the sciences.'

[89] Lodi Nauta, 'Hobbes the Pessimist?', *British Journal for the History of Philosophy* 10 (2002) 31-54, 48-50 maintains that Hobbes in fact employs the word 'reason' in two senses: as a general term for the faculty of the mind, and to refer exclusively to scientific reasoning relying on the use of words. According to Nauta, Hobbes actually employs the term 'reason' in the first sense when he maintains that natural law is accessible to reason. Nauta supports this reading by noting that while Hobbes thinks that natural law may be understood by all human beings, Hobbes also envisions that only a small minority of individuals is capable of scientific reasoning. Apart from references to the *Elements of Law*, Nauta supports this interpretation by pointing to the passage in *Leviathan* XIII.2 p.74: 'And as to the faculties of the mind – setting aside the arts grounded upon words, and especially that skill of proceeding upon general and infallible rules called science (which very few have, and but in few things), as being not a native faculty (born with us), nor attained (as prudence) while we look after somewhat else – I find yet a greater equality amongst men than that of strength.' This passage indeed indicates that Hobbes thinks that the acquired skill of reasoning grounded upon words is not common among humans. Accordingly, humans are generally equal to each other in terms of their cognitive abilities. Yet the continuation of the passage indicates clearly that Hobbes does not think that this equality in cognition is the consequence of parity in some shared natural faculty of the mind, but rather because 'prudence is but experience, which equal time equally bestows on all men [...].' First setting apart reason as a faculty uncommon among individuals, the remainder of chapter XIII then describes what happens if humans rely solely on prudence. Thus Hobbes maintains that the intellectual and physical equality of humans in the state of nature leads first to a state of diffidence and ultimately to a state of war. Only at the ending of chapter XIII does Hobbes again refer to reason, when he claims that the natural laws that follow from reason can resolve this dire situation. Read in its entirety, chapter XIII of the *Leviathan* thus hardly suggests that humans have a natural proclivity to comprehend natural law, and rather that the state of nature can be escaped only by transcending the scope of mere prudence by employing reason. The passage cited by Nauta may then suggest that regardless of Hobbes's supposition that natural law is accessible to all human beings, Hobbes is in fact not at all optimistic that all individuals can in fact be trusted to dedicate themselves to the rational deduction of natural law.

Finally the interpretations by Gert and Darwall assume that Hobbes would not have violated Hume's prohibition against deriving normative conclusions from empirical facts. Yet there is no reason to think that Hobbes actually adhered to the is/ought distinction. Apart from the fact that Hume first formulated his famous law over a hundred years after the publication of *Leviathan*, there is no solid evidence that Hobbes considered it problematic to derive values from facts. On the contrary, the case could be made that Hobbes in fact made no clear distinction between empirical facts and normative positions. Thus Hobbes applies no strict demarcations between different forms of inquiry, as he presents all forms of science, including ethics, as part of a single body of knowledge. In addition, as Noel Malcolm has noted, Hobbes did not yet accept a dichotomy between analytic and synthetic propositions later introduced by Locke.[90] Unlike Locke, Hobbes therefore did not maintain that moral concepts exist independent from the facts of empirical reality. Our own interpretation of Hobbes's moral philosophy will in fact attest that Hobbes did not make a clear distinction between facts and values. Thus we will show that he attempts to derive his conception of natural law from a conception of human nature. As this conception of human nature is derived from empirical observation, Hobbes's approach was in fact precisely to derive normative conclusions from empirical facts.[91]

Natural Law as the Means to Ends?

All the interpretations discussed above assume that for Hobbes, reasoning on the basis of language is a purely descriptive activity not concerned with the day-to-day deliberation of actions. By reasoning, we may gain a scientific understanding

[90] Malcolm, 'Hobbes's Science of Politics and his Theory of Science', 152-153 notes that Hobbes desired to provide a unified theory of knowledge with a single conception of truth. According to Malcolm, Hobbes therefore did not make the distinction later introduced by Locke between analytic propositions – which are based on conventions – and knowledge of causes derived from experience. Incidentally, this is also the reason why Hobbes appeared to present both analytical and empirical truths as conventional. Gregory Kavka, *Hobbesian Moral and Political Theory* (Princeton 1987) 6-8 makes a similar observation.

[91] Thomas A. Spragens Jr., *The Politics of Motion: The World of Thomas Hobbes* (London 1973) 179-180 comes to a similar conclusion, as he maintains that the practice of deriving the normative conclusions of natural law from the empirical fact that man desires self-preservation is already a violation of the is/ought dichotomy.

of the world, as well as deduce the universal morality of natural law. But according to the interpretations reviewed above, reason is not involved in practical deliberation. In the following, we will take a closer look at Hobbes's conception of practical deliberation and investigate its relation to morality. We will then see that reason in fact plays a crucial role in deliberating our actions and forming the will. Reason is thus not merely a theoretical, but also a practical faculty. It is this role of reason in practical deliberation that allows it to formulate the normative conclusions of natural law. Yet the precise role of reason in practical deliberation is again subject of scholarly debate. Many interpreters consider Hobbesian reason solely as an instrumental faculty that finds the means to satisfy whatever desires we might have. In the following, we will see that while reason is certainly instrumental in finding the means to ends, this reading is too simplistic by itself, and that reason is also involved in the selection of desires worthwhile pursuing. This analysis of the role of reason in deliberation will subsequently allow us to understand how reasoning dependent on language may produce the normative conclusions of natural law.

The fundamental assumption of Hobbes's account of deliberation is that the motivation for all human actions ultimately derives from the passions. Hobbes thereby rejects the notion – common to both Stoic and many scholastic moral theories – that reason may determine the will independent from the passions.[92] Instead, practical deliberation always starts with a desire or aversion to a certain object. By themselves, these desires and aversions are morally neutral, as Hobbes states that 'the desires and other passions of man are in themselves no sin.'[93] In practice, this means that while the mind is continuously affected by desires and aversions, no single passion inherently has normative priority over others. The decision which of the passions should be acted upon, and which forms of *endeavour* will have to be ignored, is therefore taken by the mind through a process of practical deliberation. Unfortunately, Hobbes's own description of deliberation, and the role of reason in this process, is often misconstrued. Hobbes describes deliberation as an 'alternate succession of appetites, aversions, hopes and fears,' thereby giving the impressions that deliberation is little more than a rotation or contest between desires and

[92] Hobbes, *Leviathan* VI.53 p.33. Cf. *De cive* I.iv.1 ann.
[93] Hobbes, *Leviathan* XIII.10 p.77.

aversions.[94] Accordingly, some interpreters have assumed that for Hobbes, human beings are wholly determined by their passions, as reason plays solely an instrumental role in the deliberation of the will. Reason would then identify the means to our desires, but is entirely dissolved from the process determining which end is worth pursuing in the first place.[95]

Consequently, these interpreters portray the role of Hobbesian reason in practical deliberation as a purely instrumental. Again, this reading is strengthened by some of Hobbes's own comments that appear to describe reason merely as a faculty finding the *means* to the *ends* determined by our desires. Thus Hobbes states that 'the thoughts are to the desires as scouts and spies, to range abroad and find the way to the things desired.'[96] Some scholars therefore seem to view Hobbes's position on the role of reason in deliberation as an anticipation of Hume's statement that 'reason is and ought to be the slave of the passions.'[97] To be sure, Hobbes thinks that reason developed by means of language is not essential to finding the means to our ends. Like animals, which possess neither language nor reason, human beings may seek to satisfy their desires by using prudence.[98] Yet as prudence is entirely dependent upon our own experience, it is both fallible and limited in scope. Hobbes thinks that by using reason, with its ability to formulate and comprehend universal propositions, we gain a much improved understanding of causal relations, which then greatly enhances our abilities to find the means to our ends.

Closely related to this interpretation of reason as solely instrumental in the achievements of our desires, is the view that Hobbesian natural law is

---

[94] Hobbes, *Leviathan* VI.51 p.33. Schneewind, *The Invention of Autonomy* 88 thus describes deliberation merely as a 'conflict of desires.'

[95] Sorell, *Hobbes* 95; Darwall, 'Normativity and Projection in Hobbes'ss Leviathan' 331n31; Patrick Riley, *Will and Political Legitimacy: A Critical Exposition of Social Contract Theory in Hobbes, Locke, Rousseau, Kant, and Hegel* (Cambridge MA 1982) 37-38; Daniela Coli, 'Hobbes's Revolution' in: Victoria Kahn et al., *Politics and the Passions 1500-1850* (Princeton 2006)73-92, 82-83; Malcolm, 'Hobbes and Spinoza', in: idem, *Aspects of Hobbes* 27-52, 30-31; Schneewind, *The Invention of Autonomy* 88.

[96] Hobbes, *Leviathan* VIII.16 p.41. Cf. III.4 p.13: 'From desire ariseth the thought of some means we have seen produce the like of that which we aim at; and from the thought of that, the thought of means to that mean; and so continually, till we come to some beginning within our power.'

[97] Coli, 'Hobbes's Revolution' 83; Cf. Hume, *A Treatise of Human Nature* II.iii.3 for the passage. Although in isolation this quote also misrepresents Hume's moral philosophy, as Hume did not think that human actions follow exclusively from selfish passions – they may also follow from moral sentiments.

[98] Hobbes, *Leviathan* III.7 p14.

nothing more than a set of prudential guidelines indicating the most effective *means* to felicity, conceived as the continuous satisfaction of whatever desires we might have. According to this reading, the pursuit of self-preservation decreed by the law of nature is the *means* towards the satisfaction of all other desires, simply because being alive is the necessary prerequisite to the satisfaction of any desire.[99] The laws of nature themselves are then the means to best realise this pursuit of self-preservation. Hobbes therefore states that 'it is a precept, or general rule, of reason *that every man ought to endeavour peace*,' because peace is the most effective means to self-preservation, which in turn is a necessary means to the satisfaction of our desires.[100] Following this interpretation, Hobbes does not consider the pursuit of self-preservation as imperative because it follows from a desire that should be considered inherently moral. As desires are by themselves morally neutral, Hobbes maintains that the pursuit of self-preservation is rational merely because it is a necessary *means* to all other ends.

Practical Deliberation

This interpretation of reason as instrumental to achieving our ends certainly describes one of the more important roles of reason in practical deliberation. As we have seen above, the cognitive functions of language allow human beings to extend their understanding of causes beyond their own particular experiences, thereby supplementing mere prudence with the faculty of reason. This artificial faculty then allows human beings to formulate purely descriptive scientific laws. But there is no ground to the assumption that this improved understanding of consequences cannot also greatly expand our capacities for finding the means to objects desired. John Deigh's position that solely prudence, and not reason, is used to find the means to ends is therefore unfounded.[101] In fact, following the lead of van Apeldoorn and Abizadeh, a closer reading of Hobbes's account of practical deliberation will demonstrate that reason has a role in the

[99] Haakonssen, *Natural Law and Moral Philosophy* 31.

[100] Hobbes, *Leviathan* XIV.4 p.80.

[101] Deigh, 'Reason and Ethics in Hobbes's *Leviathan*' restricts reason to nothing more than formulating scientific conclusions from the conventional meaning of words. Even apart from the objections to his interpretation of Hobbes's conception of reason voiced above, Deigh's view that reason is not involved in finding the means to ends is not only incomprehensible, it also contradicts Hobbes's own statements.

determination of our actions even more extensive than supposed by the commentators presenting reason as purely instrumental.[102]

We should begin this reading by acquiring a more thorough understanding of Hobbes's conception of desire and aversion. In line with his materialist metaphysics, Hobbes thinks that like all other cognitive processes, these desires and aversions ultimately depend on material processes in the body. He maintains that desires and aversions originate in the imagination, where memories of pleasure and pain associated with certain objects give rise to voluntary motion or *endeavour* towards these objects of experience. Yet while they are ultimately based in material processes, most desires and aversions are not simply the consequence of non-cognitive and corporeal urges. Apart from hunger, thirst and some other basic appetites, Hobbes thinks 'appetites of particular things, proceed from experience and trial of their effects upon themselves or other men.'[103] The passions therefore arise in the imagination and depend on experience, as the memory of pleasurable or evil consequences of attaining a certain object in the past will give rise to a desire to reclaim or avoid the same object in the future. For Hobbes, the passions are therefore clearly cognitive phenomena that depend on ideas of objects and their consequences in the imagination.[104]

In line with this more precise understanding of Hobbes's conception of the passions, we may now revise the interpretation of practical deliberation encountered in the previous section. A closer reading of Hobbes's account of deliberation reveals that the will does not simply follow from blind passion, but that reason plays a significant role as well:

> When in the mind of man appetites and aversions, hopes and fears, concerning one and the same thing arise alternately, and diverse good and evil consequences of the doing or omitting the thing propounded come successively into our thoughts, so that sometimes we have an appetite to it, sometimes and aversion from it, sometimes hope to be able to do it, sometimes despair or fear to attempt it, the whole sum of

[102] Abizadeh, 'Hobbes on Mind'; van Apeldoorn, 'Reconsidering Hobbes's Account of Practical Deliberation.'

[103] Hobbes, *Leviathan* VI.4 p.28. A.P. Martinich, *Hobbes* (New York and London 2005) 38-39 appears to have missed this particular point, as he objects that Hobbes's characterization of desires is incorrect, as desires like hunger and thirst follow directly from corporeal processes rather than movements in the imagination.

[104] Hobbes, *Leviathan* VI.1-2 p.27-28; *De corpore* IV.xxv.12 p.406-408. Cf. James, *Passion and Action* 129.

> desires, aversions, hopes and fears, continued till the thing be either done or thought impossible, is that we call *deliberation*.[105]

In this passage, Hobbes indicates first of all that deliberation is not merely an unregulated alteration or contest of desires producing the will. When deliberating, the mind rather examines all the possible consequences of attaining an object of desire, as well as of the actions required to achieve this end. Hobbes concedes that we do not always have time for careful deliberation, and therefore often act on a whim.[106] But when we have the time and inclination to deliberate, the mind then weighs the foreseen good against potential evil consequences and then decides whether an object should be pursued or not. We could therefore see deliberation as the process of deciding which desires should be acted upon, and which should be ignored.[107] Yet it is probably even more accurate to describe deliberation as the mental discourse by which we decide which objects we have reason to desire – in other words, which objects reason considers desirable in the first place.[108]

Apart from a limited number of purely bodily desires, the passions are therefore not merely mindless urges that incite human actions – as some interpretations have supposed.[109] Instead, they are more akin to beliefs that

---

[105] Hobbes, *Leviathan* VI.49 p.33. Cf. VI.57 p.34; *De Corpore* XXV.13 p.408; *Questions concerning Liberty, Necessity and Chance* XXVI ed. William Molesworth (London 1841). p.357-358: 'Secondly, I conceive when a man deliberates whether he shall do a thing or not do a thing, that he does nothing else but consider whether it be better for himself to do it or not do it. And to consider an action, is to imagine the consequences of it, both good and evil. From whence is to be inferred, that deliberation is nothing but alternate imagination of the good and evil sequels of an action.'

[106] Hobbes, *Questions concerning Liberty, Necessity and Chance* XXV p.344: 'I conceive that when it cometh into a man's mind to do or not do some certain action, if he have no time to deliberate, the doing or abstaining necessarily followeth the present thought he had of the good or evil consequences thereof to himself.'

[107] Gauthier, *The Logic of the Leviathan* 12 provides a similar reading of the role of reason in deliberation.

[108] In line with his mechanistic materialism, Hobbes does think that this process of deliberation is causally determined, as human beings do not possess freedom of the will. Yet Hobbes's determinism need not interfere with our interpretation. Our reading does not conclude that reason liberates the will from causality. Rather, our interpretation simply claims that reason, however causally determined, participates in the determination of the will. Cf. Hobbes, *Questions concerning Liberty, Necessity and Chance* XXX p.372-373.

[109] Schneewind, *The Invention of Autonomy* 84-85; MacIntyre, *A Short History of Ethics* 134: 'Hobbes's determinism [...] leads him to treat our desires as given and unalterable. The criticism of our desires and their rational remolding have no place in the Hobbesian system.' A similar point is

objects or actions will produce pleasure and avoid pain. In practice, these beliefs are often short-sighted or mistaken, leading us to imminent pleasure but unfortunate long-term consequences. Yet we may improve our chances for lasting felicity by careful deliberation on the potential consequences of desired objects or actions. During this process of deliberation, the mind will typically consider a whole range of beliefs about objects and their possible consequences. These beliefs do not arise spontaneously, nor are they purely determined by our physical constitution. As ideas in the imagination, desires are part of our mental discourse, and occur in reaction to preceding reflections – including those of practical deliberation. During this succession of beliefs about the consequences of actions or objects, the mind will reject all desires deemed unfavourable, and finally act upon the one passion that promises the most durable form of felicity. Consequently, desires and aversions are both the object and the product of active deliberation in which the mind judges the good or evil consequences of pursuing or avoiding objects in the imagination.

Finally, this reading is further supported by Hobbes's own explicit comparison between the use of reason in practical deliberation leading to the will, and the use of reason in search of scientific truths producing 'judgment.' Hobbes thus describes the process of reasoning about empirical objects as the alternation of 'opinions' – thoughts or hypotheses about the consequences of a certain object. During this mental discourse, the mind considers various opinions about the object imagined, using experience and/or reason to determine their validity. This process ends only when all but one opinion have been discarded as unfounded. Hobbes therefore states that 'the last opinion in search of truth of past and future is called the *Judgment*.'[110] Hobbes subsequently indicates that practical deliberation in fact proceeds through an analogous process, in which deliberated but restrained desires compare to opinions and the will to judgment.[111] Thus in practical deliberation the mind experiences various desires and aversions by considering the possible good or evil consequences of an object or action. Like in the case of judgment of opinions, this deliberation ends when

made by Michael Allen Gillespie, *The Theological Origins of Modernity* 236 who also maintains that reason solely finds the means to our passions, but does not direct or control them.

[110] Hobbes, *Leviathan* VII.2 p.35.

[111] Hobbes in fact calls these restrained desires intentions or inclinations in the *Questions concerning Liberty, Necessity and Chance* XXVII p.360: 'All other appetites to do and to quit, that come upon a man during his deliberation, are usually called intentions and inclinations, but not wills.'

all but one desire have been rejected as ultimately insufficiently favourable in its consequences. Hobbes therefore describes the will as nothing but 'the last appetite in deliberating.'[112] Hobbes's own comparison between the process of reasoning about empirical objects and that of practical deliberation then provides additional – albeit circumstantial – evidence for our reading that reason is involved in the determination of the will.

Normative Judgment

Following this reappraisal of Hobbes's account of desire and deliberation, we may now distinguish more precisely the role of reason in both determining our actions and producing normative conclusions. We have seen that practical deliberation is not just a blind alteration of desires, but a cognitive process in which the pleasurable or painful consequences of an object or action are weighed, ultimately giving rise to the will. As with purely descriptive mental discourse producing judgment, deliberation thus relies on an understanding of causal relations. We have seen in the first section that the understanding of the consequences of objects or actions depends in turn on either prudence derived from experience, or the ability to reason acquired with language. While deliberating with the help of prudence, both man and beast use their personal experience of particular objects to decide whether these objects and their consequences are desirable or not. Yet as we have seen, Hobbes thinks that prudence may only generate a limited understanding of causal relations. Only by using reason do we gain the ability to perceive and understand long chains of consequences attached to certain objects or actions. By learning a language and acquiring the ability for reasoning, human beings – as opposed to animals – therefore not only increase their understanding of causal relations in nature, giving rise to superior judgment. They also significantly enhance their ability to foresee the pleasurable or painful consequences of objects or actions.

Yet as practical deliberation is used to formulate the will, the same process is also responsible for normative judgment on the level of the individual. Hobbes can be classified as a moral subjectivist. This entails that he denies that there is some moral rule to be taken from the objects themselves, and that individuals diverge in their normative evaluation of objects and actions.[113]

[112] Hobbes, *Leviathan* VI.53 p.33.
[113] Hobbes, *Leviathan* VI.7 p.29.

Accordingly, Hobbes maintains that 'whatsoever is the object of any man's appetite or desire that is it which he for his part calleth *good*; and the object of his hate and aversion; *evil*.'[114] This passage may seem puzzling at first, as it appears to suggest that our normative judgment is somehow determined by our passions. Yet in light of our interpretation of Hobbes's account of practical deliberation, it is clear that this passage does not entail that our estimation of good and evil is simply the consequence of whatever passions we might have at any given moment. Instead, we have seen that passions are more akin to beliefs about the pleasurable of painful consequences of objects. And ideally we consider carefully the prospective consequences of objects or actions before acting upon the last desire of our deliberation, thereby producing the will.[115] In short, when we call an object good or evil, this simply entails that we foresee that it will have either pleasurable or painful consequences. On the subjective level, methodical normative judgment is therefore nothing other than a careful deliberation about consequences of an object or action. Accordingly, we could classify Hobbes's conception of normativity as *consequentialist*.

This interpretation of Hobbes's conception of morality as consequentialist is corroborated by a brief philological digression in the *Leviathan* on the various meanings of the words 'good' and 'evil.' Hobbes explains that terms good and evil accommodate multiple significations not easily distinguished in English, but differentiated in Latin by three separate terms. Thus the word *jucundum* refers to the pleasurable effect of a certain object, while *utile* signifies that a certain object is good because it is a valuable means towards satisfying our desires. Furthermore, Hobbes points out that the Latin *pulchrum* is used to denote: 'that which by some apparent signs promiseth good.'[116] The word 'evil' can be translated by a similar triad of terms: *molestum* for evil in effect, *inutile* for evil in the means, and *turpe* for the promise of evil. This digression then illustrates that Hobbes thinks that we do not consider an object good or evil without reason, but rather on the basis of our foresight of the consequences it has for us. Thus Hobbes's discussion of the term *pulchrum* indicates that when we desire an object and therefore call it good, this signifies that we have reason to conclude that an object will have favourable consequences in the future. This

[114] Hobbes, *Leviathan* VI.7 p.28. Cf. *De Homine* VII.3 p.32.

[115] Thus in *De cive* I.2 p.43 Hobbes describes the will as 'that, which every one of those, who gather together, propounds for himself for good.'

[116] Hobbes, *Leviathan* VI.8 p.29

again indicates that for Hobbes, normative judgment of objects or actions follows from the prediction of potential pleasurable or painful consequences in future.

In the remainder of this chapter, we will explore how our reading of Hobbes's account of subjective normative judgment as consequentialist affects the interpretation of his theory of morality and natural law. For now, we should note that following our consequentialist reading, Hobbes's conception of normative judgment could at the same time be described as a species of moral intellectualism. We have seen that our capacity for normative judgment – and thereby our ability to identify and pursue our long-term interests – is dependent on our respective ability to foresee relations of cause and effect in nature. As reason is an integral part of deliberation and normative judgment, any failure to identify and pursue our interests therefore cannot simply be attributed to misdirected passions. Rather, these failures follow from either defective or careless reasoning, failing to foresee the consequences of our desires and the means to achieve them. The corollary of this moral intellectualism is then that no individual does wrong willingly. In Hobbes's case, this means that no person would voluntarily and knowingly harm her self-interest, but only following misguided reasoning about the consequences of actions.

Reason and Sociability

On the basis of this consequentialist and intellectualist understanding of Hobbes's conception of practical deliberation and subjective normative judgment, we should now consider the question to what extent reason allows humans to achieve their ends and live together in society. In theory, normative judgment may be perfectly accurate as long as we deliberate carefully about all the prospective consequences of our actions. Yet if human beings would indeed have the ability to perceive their long-term interests through the rational understanding of the consequences of their actions, this raises the question why Hobbes is so pessimistic about the natural condition of mankind. If human beings have such a potent instrument for perceiving the consequences of their actions, why will they not naturally live in harmony among each other? In this respect, Hobbes departed significantly from both Aristotle and his immediate predecessor Grotius, who had argued that it was precisely our rational faculty that allowed humans to become social animals.[117] By contrast, Hobbes maintains that some

species of animals lacking reason – including bees and ants – are indeed naturally social, but that human beings do not display sociability as a natural tendency.[118] In the first instance, human sociability is therefore not natural because it relies on the artificial faculty of reason. Yet Hobbes also thinks that by itself, the capacity for reasoning is not sufficient to transform human beings into truly sociable creatures.

The most immediate reason why Hobbes thinks that human beings are not naturally sociable can be found in the fact that the normative judgment of individuals tends to diverge considerably. Hobbes thinks that this disagreement in normative judgment has two underlying causes in practical deliberation:

> And therefore the voluntary actions and inclinations of all men tend, not only to the procuring, but also to the assuring of a contented life, and differ only in the way; which ariseth partly from the diversity of passions in divers men, and partly from the difference of the knowledge or opinion each one has of the causes which produce the effect desired.[119]

Hobbes's thus maintains that while all human beings strive for felicity, they disagree both about what felicity entails concretely, and about the best means to achieve it. In the first instance, this disagreement is caused by a diversity in the passions, which is in turn the consequence of diverging predilections, as humans consider different consequences pleasurable or not. In line with his materialist ontology, Hobbes thinks that these variations in taste depend partially on education, and partially on corporeal factors, as differences in physical constitution and fluctuating bodily needs may render the same object pleasurable to some while abhorrent to others.[120] Even if human beings would have an infallible understanding of causality, these differences in predilections would nonetheless ensure disagreement in subjective normative judgment. As variations in their physical constitution cause individuals to disagree about the things they find desirable, they will never be able to formulate a *finis ultimus*

---

[117] Grotius, *On the Law of War and Peace* ed by Stephen Neff (Cambridge 2012) p.3-4; Aristotle, *Politics* I.ii.1253a1.

[118] Hobbes, *Leviathan* XVII.6 p.108; *De cive* I.2 p.42.

[119] Hobbes, *Leviathan* XI.1 p.57-58.

[120] Hobbes, *Leviathan* VIII.14 p.40

that indicates the universal end of all human beings.[121] Hobbes thereby rejects the possibility of a teleological morality of the Aristotelian kind.[122]

Unlike naturally social animals, human beings thus have different desires, and therefore diverging interests.[123] This property of human nature is sometimes considered the main cause for the conflicts of the state of nature, and consequently the primary reason why human beings are not naturally social. But this analysis is not an exhaustive explanation why humans are not naturally social beings. As Hobbes states in the passage quoted above, individuals also differ in their comprehension of causal relations, leading to dissenting opinions on how to achieve their ends. These differences will be particularly prevalent when human beings employ prudence – derived solely from their own personal experience – in deliberation, thereby exacerbating their normative disagreement even further. Yet as we have seen, Hobbes thinks humans are also capable of using reason in deliberation to better foresee the remote consequences of their actions. Presumably, this would enable them to conclude that it is in their collective interest to set aside their differences and see beyond their own particular views on good and evil. Accordingly, individuals would pursue their own desires to the extent that they do not harm the interests of others, thereby evading conflict and enabling them to live in harmony among each other. In other words, despite their conflicting views on good and evil, reason would allow human beings to become sociable after all.

On the one hand, Hobbes agrees that this is indeed possible, as he states that despite the normative disagreement among individuals, 'reason suggesteth convenient articles of peace, upon which men may be drawn to agreement. These articles are they which otherwise are called the Laws of Nature.'[124] These natural laws therefore collectively guide individuals towards a more harmonious existence in which the most violent conflicts of interest are neutralised, while individuals remain free to pursue their desires to the extent that they do not

[121] An additional reason why Hobbes thinks that man does not have a *finis ultimus* is that humans will never be able to satisfy or extinguish all of their passions, as interaction with their environment will always give rise to new desires. As humans will never be able to find repose from their passions, any search for tranquillity of mind would be in vain. See Hobbes, *Leviathan* VI.58 p.34-35.

[122] Hobbes, *Leviathan* XI.1 p.57.

[123] Hobbes, *Leviathan* XVII.8 p.108 suggests that social animals like bees and ants are naturally social because they do not experience conflicts of interest.

[124] Hobbes, *Leviathan* XIII.14 p.78.

harm others. Thus while Hobbes denies that human beings are naturally social, reason nonetheless enables us to perceive which kind of actions we should perform or forbear in order to end the conflicts of the state of nature and live in society with our neighbours.[125] On the other hand, Hobbes is not at all optimistic that human beings will follow the laws of nature on their own initiative. Consequently, resolving the conflicts of the state of nature requires the erection of an all-powerful sovereign to keep all subjects in awe.[126] This then raises the question why Hobbes thinks that despite man's potential rationality, human beings nonetheless tend to disregard the rational guidelines of natural law.

In line with our intellectualist reading, Hobbes's pessimism – or realism – in this respect may be explained as the consequence of the shortcomings of reason. We have already seen that Hobbes concedes that it is not realistic to assume that human beings will always use their reason when deliberating – often they will just act on a whim. Yet even if human beings would always take care to reason about their actions, the erection of an all-powerful sovereign would still be necessary because human reason will never be infallible.[127] Thus for all its perspicuity, reason will never allow us to foresee all possible consequences with complete accuracy.[128] Furthermore, the fact that reason is an acquired rather than an innate faculty means that the ability to effectively apply reason in deliberation is dependent upon a hardly commonplace level of education. In addition, Hobbes concedes that even the mental reckoning of the most able reasoners is not immune to errors of computation.[129] Finally, Hobbes recognises that while language is the medium that enables our reasoning, this same

[125] Hobbes, *De cive* ed. Howard Warrender (Oxford 1983) II.1 ann p.52: 'By Right reason in the natural state of men, I understand not, as many doe, an infallible faculty, but the act of reasoning, that is, the peculiar and true ratiocination of every man concerning those actions of his which may either redound to the dammage, or benefit of his neighbours.'

[126] Hobbes, *Leviathan* XIII.4 p.75: 'men have no pleasure, but on the contrary a great deal of grief, in keeping company where there is no power able to over-awe them all.'

[127] Hobbes, *Leviathan* V.3 p.23: 'no one man's reason, nor the reason of any number of men, makes the certainty [...] And therefore, as when there is a controversy in an account, the parties must by their own accord set up for right reason the reason of some arbitrator or judge whose sentence they will both stand, or their controversy must either come to blows of be undecided.'

[128] Hobbes, *Leviathan* VI.57 p.34: 'And because in deliberation the appetites and aversions are raised by foresight of the good and evil consequences and sequels of the action whereof we deliberate, the good and evil effect thereof dependeth on the foresight of a long chain of consequences, of which very seldom any man is able to see to the end.'

[129] Hobbes, *Leviathan* V.3 p.23: 'in any other subject of reasoning, the ablest, most attentive, and most practiced men may deceive themselves and infer false conclusions.'

instrument may also be used to lead us astray. Above we have already seen that Hobbes is concerned that metaphorical, absurd or rhetorical uses of language have the ability to sway our opinions regardless of the truth. The same uses of language may also be used to influence views about our interests.[130] In other words, language greatly improves our capacity to foresee the good or evil consequences of things, but it also has the ability to deceive our judgment. This is one of the reasons why Hobbes thinks that the sovereign should have the right to police the opinions disseminated among its subjects.[131]

These factors combined make that reason is far from an infallible guide in normative judgment or the pursuit of felicity. This would be problematic for individuals subsisting by themselves, but it is an even more serious issue for human beings living together. Not only do we lack certainty in our own normative judgment. We are also faced with the fact that in society, we cannot rely on others to foresee the consequences of their actions. This is especially precarious if the interests of these others conflict or overlap with those of ourselves. Hobbes illustrates this problem by positing a 'fool, who hath said in his heart that there is no justice.'[132] This fool has concluded that in certain situations, it would have beneficial consequences if he would break covenants made – thereby abusing the interests of others and breaking the third law of nature. Hobbes states that while this course of action may seem attractive to some, 'this specious reasoning is nonetheless false.'[133]

Although there is considerable scholarly debate on the precise interpretation of the reply to the fool, Hobbes appears to be saying that the fool does not realise that his transgression is likely to provoke retribution – both in society and the state of nature.[134] The fool is thus not someone who cannot help

[130] Hobbes, *Leviathan* VIII.27 p.46-47; IV.13 p.19: 'as men abound in copiousness of language, so they become more wise, or more mad, than ordinary.' Cf. Quentin Skinner, *Reason and Rhetoric in the Philosophy of Hobbes* (Cambridge 1997) 347 who has argued that Hobbes had great confidence in the abilities of reason in the beginning of his career, but became more pessimistic in the *Leviathan* as he became more sensitive to the power of rhetoric to sway opinion. For a critique of Skinner's argument, see Nauta, 'Hobbes the Pessimist?'.

[131] Hobbes, *Leviathan* XVIII.9 p.113.

[132] Hobbes, *Leviathan* XV.4 p.90.

[133] Hobbes, *Leviathan* XV.4 p.91. Cf. *De cive* II.1 ann. p.52-53: 'the whole breach of the Lawes of Nature consists in the false reasoning, or rather folly of those men who see not those duties they are necessarily to performe to others in order to their owne conservation.'

[134] Kavka, *Hobbesian Moral and Political Theory* 137-156 agrees and provides an extensive analysis of the issue and concludes that overall, the fool would be advised to keep his covenants in

himself; whose action is caused by some overwhelming desire he cannot ignore. In addition, the fool does not deny that there is such a thing as justice. He knows that the breaking of covenant is a violation of a rule that would contribute to collective felicity if all agreed to adhere to it. His transgression is therefore not simply the product of moral depravity. Rather, the fool's mistake consists in a failure to foresee that the temporary gain of his transgression will be offset by more remote negative consequences. In other words, the fool's deliberation has made use of spurious or short-sighted reasoning.[135]

Hobbes's discussion of the fool indicates that the primary cause of the conflict of the state of nature is neither the inconsistency of the desires of individuals, nor the supposed subservience of reason to mindless passions. Rather, the ultimate cause of these conflicts is to be found in the shortcomings of human reason in practical deliberation. Hobbes's fool illustrates that individuals cannot be counted on to accurately foresee the consequences of their actions and act in their own long-term interest. Consequently, persons in the state of nature, where conduct is not regulated by written laws enforced by a sovereign, can never be certain that the individuals they meet will deliberate to act according to the laws of nature. This is then at least one of the underlying reasons why Hobbes describes relations between individuals in the state of nature as one of diffidence of one another.[136] And in this state of diffidence, it will often be in our immediate self-interest to act in ways that harm the interests of others, thereby also endangering our own in the long. Due to man's relative equality, this

---

any situation. Darwall, *The British Moralists and the Internal 'Ought'* 75-76 maintains that Hobbes sees the violation of covenants as irrational because it could lead to the expulsion of the agent from civil society, exposing him to the inherently worse situation of the state of nature. Michael LeBuffe, 'Hobbes's Reply to the Fool' *Philosophy Compass* 2 (2007) 31-45 disagrees and finds Hobbes's reply unconvincing.

[135] Michael Byron, 'Hobbes's Confounding Foole', in: S.A. Lloyd (ed.), *Interpreting Hobbes's Political Philosophy* (Cambridge 2019) 206-222 agrees that the reasoning of the fool is spurious, both in a conceptual sense, as well as in its outcomes. S.A. Lloyd, *Morality in the Philosophy of Thomas Hobbes* (Cambridge 2009) 305-307 roughly agrees, pointing to Hobbes's Latin revision of the passage on the fool, which adds a passage that indicates that any benefit the fool would receive from acting unjustly would happen by chance, and could not be clearly foreseen. Injustice would therefore pay only when the fool has the luck not to be caught.

[136] Hobbes, *Leviathan* XIII.3-4 p.75. To be sure, Hobbes thinks that this diffidence is also the consequence of the relative equality in physical and mental capacities, inviting attempts by individuals to overpower each other by force or cunning. Yet this factor would never be relevant but for the insufficiency of subjective reasoning to reconcile the desires of individuals by ensuring the collective obedience to the law of nature.

dynamic inevitably leads to a state of war. For Hobbes, the only way to overcome this situation is then by denying subjects the right to normative judgment on all but purely individual matters, and entrusting an all-powerful sovereign with both the interpretation and enforcement of justice.

Natural Law

We may now return to the question posed in at the beginning of this chapter, and explain how Hobbes thinks reason may deduce natural law. In the preceding sections, we have already seen that language greatly increases our understanding of causal relations in nature, thereby producing the capacity for reasoning. Subsequently, we discovered that reason plays an important role in practical deliberation, as it allows us to foresee the consequences of both desires and the means to satisfy them. By the same method, reason also participates in normative judgment – the subjective evaluation of the good or evil consequences of objects of actions. This normative judgment is self-interested – although not necessarily selfish – as it seeks to attain pleasurable and avoid painful consequences.[137] Yet we have also seen that this normative judgment is both fallible in predicting the consequences of actions, and varies among individuals due to diverging predilections and differences in experience and education. These properties of subjective normative judgment are the fundamental reason why human beings are not naturally social, and often come into conflict with each other.

In response, Hobbes outlines a set of natural laws that prescribe how human beings may live together despite their conflicts of interest and fallible judgment of consequences. These natural laws prescribe how human beings may best pursue their self-interest in the company of others who seek to do the same. As in the case of subjective normative judgment, the reasoning producing natural law relies on a grasp of the consequences found in nature. Yet because natural laws do not merely provide moral guidance for solitary beings, but rather for life in society, the rational deduction of natural law also requires an understanding of the causal relations governing the behaviour of the human beings themselves. This is then the reason why Hobbes's deduction of natural law follows from a detailed investigation of human nature, outlining both man's intellectual

[137] Kavka, *Hobbesian Moral and Political Theory* 44-50 explains how Hobbes's egoism is self-interested, but not necessarily self-centered.

capacities and his overall needs and desires. On the basis of this understanding of human nature, Hobbes first identifies which kinds of desires are most amenable to human society, and secondly the means by which these desires can be satisfied most effectively. The role of reason in deducing natural law is therefore similar to its function in practical deliberation – it weighs desires in view of their consequences, and finds the means to the things desired.[138]

In Hobbes's own deduction of natural law, this scrutiny of the various desires and fears commonly found in human beings occurs in chapter XI of the *Leviathan*, just before the description of the state of nature. Hobbes begins by stating that humans not only have diverging passions, but also that they will never find complete repose from them. Subsequently, he maintains that these various passions incline humans to desire power in order to attain their various desires. Hobbes then describes how the resulting 'competition of riches, honour, command, or other power, inclineth to contention, enmity, and war.'[139] According to Hobbes, this desire for power is thus one of the main causes of the conflicts that characterise the state of nature. Yet Hobbes also describes which kinds of desires are liable to incline human beings towards a sociable life in obedience to the sovereign:

> Desire of ease and sensual delight disposeth men to obey a common power, because by such desires a man doth abandon the protection might be hoped for from his own industry and labour. Fear of death and wounds disposeth to the same, and for the same reason.[140]

Although Hobbes states elsewhere that 'the desires and other passions of man are in themselves no sin,' this apparently does not entail that we should not judge desires by their consequences.[141] Hobbes here clearly identifies which passions contribute to the conflicts of the state of nature, and are therefore ultimately undesirable, and which desires cause humans to accept a social existence under the rule of a sovereign.

---

[138] This is why Hobbes, *Leviathan* XIII.13 p.78 states that: 'And thus much for the ill condition which man by mere nature is actually placed in, though with a possibility to come out of it, consisting partly in the passions, partly in his reason.'

[139] Hobbes, *Leviathan* XI.3 p.58.

[140] Hobbes, *Leviathan* XI.4 p.58. Cf. XII.14 p.78. Hobbes subsequently adds that the desire for knowledge has a similar effect on people.

[141] Hobbes, *Leviathan* XIII.10 p.77.

Subsequently, Hobbes's account of the state of nature provides a more concrete illustration of the consequences of the various passions inherent in man. Here again we find that besides diffidence, the main causes of war in the state of nature are the desire for glory as well as competition for power.[142] But after describing the horrible consequences of pursuing these desires, Hobbes then outlines the method for escaping this state of war:

> And thus much for the ill condition which man by mere nature is actually placed in, though with a possibility to come out of it, consisting partly in the passions, partly in his reason. The passions that incline men to peace are fear of death, desire of such things as are necessary to commodious living, and a hope by their industry to obtain them.[143]

This passage again suggests that on the basis of his rational investigation of human nature, Hobbes makes an implicit normative evaluation of the various passions prevalent in man. Employing a method similar to that used in practical deliberation, Hobbes has weighed the desires and aversions inherent in human nature in view of their consequences, and has concluded that the desire for commodious living, as well as fear of death, are to take precedence above most other passions – simply because these desires are ultimately more likely to have pleasurable consequences.[144]

After drawing this conclusion, Hobbes promptly continues with his deduction of the right of nature, which allows one to do anything in pursuit of self-preservation, as well as the law of nature, which forbids anything that endangers our life or the means of preserving it. The natural laws themselves are then moral guidelines that establish a peaceful settlement of the conflicts of the state of nature, thereby safeguarding the preservation all individuals living in

---

[142] Hobbes, *Leviathan* XIII.6-7 p.76.

[143] Hobbes, *Leviathan* XIII.13-14 p.78.

[144] Hobbes, *Leviathan* XIII.10 p.76-77 in fact himself uses the term 'weigh' to describe his method: 'it may seem strange, to some man that has not well weighed these things, that nature should thus dissociate, and render men apt to invade and destroy one another. And he may, therefore, not trusting to this inference made from the passions, desire perhaps to have the same confirmed by experience.' A similar reading of Hobbes is proposed by Hirschmann, *The Passions and the Interests* 31, who suggests that by means of reasoning humans may weigh their passions and determine which are most conducive to their well-being, concluding that fear of death and the desire for a commodious living are to be desired over riches, glory, and dominion. According to Hirschmann, this conclusion is what leads humans to institute the covenant erecting a sovereign that would ensure their preservation and well-being.

society. Yet in the *Leviathan*, we find no clear explanation why only the desire for self-preservation – or fear of death – is fundamental to both the right and law of nature, while other sociable desires such as for a commodious living or for knowledge, appear to be no longer relevant. Some interpreters have therefore suggested that self-preservation does not just to refer to bare survival, but also includes the ulterior pursuit of convenience.[145] Others have maintained that self-preservation should simply be conceived as the means towards the satisfaction of all other desires.[146]

Following our own interpretation of Hobbes's account of practical deliberation, we may also provide an alternative explanation for the normative primacy of self-preservation over all other desires. We have seen that due to differences in experience and inclinations, individuals generally have diverging desires and fears. Accordingly, even sociable passions like the desire for knowledge and ease of living are not universally present among human beings, as not everyone will understand the favourable consequences of pursuing these passions. Yet Hobbes thinks that everyone will be able to foresee the consequences of death, and will therefore strive for their own self-preservation.[147] This universal fear of death is then the reason why the desire for self-preservation takes normative precedence above all other ends. Consequently, man always retains the natural right to defend his life when it is threatened – simply because it is always in his interest.[148] More importantly, the fact that all human beings recognise the consequences of death is the most likely reason why Hobbes constructed his formulation of natural law upon the universally prevalent desire for self-preservation. As the need for self-preservation is evident to all, the precepts of natural law will be recognisable to all individuals who would consult their reason.

---

[145] This suggestion is found in Kavka, *Hobbesian Moral and Political Theory* 82.

[146] Haakonssen, *Natural Law and Moral Philosophy* 31.

[147] Hobbes, *De cive* I.7 p.47: 'For every man is desirous of what is good for him, and shuns what is evill, but chiefly the chiefest of natural evils, which is Death; and this he doth, by a certain impulsion of nature, no lesse then that whereby a Stone moves downward: It is therefore neither absurd, nor reprehensible; neither against the dictates of true reason for a man to use all his endeavours to preserve and defend his body.'

[148] Hobbes, *Leviathan* XVI.11-16 p.141-142. Individuals retain this right to resist a threat to their lives in society, even if they have actually transferred their natural right to do anything in their power to secure their self-preservation to their sovereign.

Conclusion

In this chapter, we have interpreted Hobbes's theory of normative judgment by reconstructing his account of the way in which human beings employ their reason to perceive morality and natural law. We have begun our discussion by investigating Hobbes's conception of reason, devoting particular attention to the role of language in the development of this faculty. Meanwhile, we have considered and ultimately dismissed the interpretation of John Deigh, who has claimed that as Hobbes thinks that reason is developed only by means of language, he must also have thought that normative reasoning producing natural law is determined by the conventions of language. Subsequently, we have discussed the interpretations of Gert and Darwall, who claim that Hobbes's account of reason by means of language solely describes the process of reasoning concerning empirical objects. As they assume that Hobbes would not have violated the distinction between facts and values, Gert and Darwall claim that Hobbes's account of reason by means of language cannot account for normative reasoning. Accordingly, they each provide an interpretation that ascribes to Hobbes the notion that normative reason proceeds from some innate faculty or property. We have not only shown that this interpretation is insufficiently supported by textual evidence, but also relies on the unsubstantiated assumption that Hobbes respected the is/ought dichotomy.

In our search for Hobbes's theory of normative reasoning, we have subsequently investigated his account of practical deliberation, which describes the mechanism by which human beings deliberate their actions and formulate their will. At the same time, we have considered the common interpretation that ascribes to Hobbes the view that all human thoughts and actions are the consequence of the passions, while the sole purpose of reason is to find the means towards the satisfaction of our desires. Accordingly, this reading considers Hobbesian natural law as nothing more than a prudential rule that indicates the most effective means towards felicity, conceived as the sustained satisfaction of whatever passions we might experience. Our own reading of Hobbes's account of practical deliberation has then shown that reason is not merely an instrumental faculty that finds means to ends, but in fact plays a crucial role in scrutinising desires and the formulation of the will by foreseeing the consequences of actions considered in deliberation. We have then shown that Hobbes's account of practical deliberation thereby also functions as an account of subjective

normative judgment, as it demonstrates the way individuals may evaluate an object or action by foreseeing its consequences.

Subsequently, we have seen that this consequentialist reading of Hobbes's theory of normative judgment also has repercussions for our interpretation of his conception of natural law. We have started our investigation with the question why Hobbes thinks that the desire for self-preservation should provide the basis for his conception of natural law. The most common explanation has been that Hobbes thinks that self-preservation is the prerequisite, and therefore the necessary means, to felicity. Yet our own interpretation has shown that reason is not merely instrumental to finding the means to ends, but also involved in determining which ends are worth pursuing. Accordingly, we have explained that Hobbes did not consider self-preservation as the goal of natural law merely because it is a means to felicity. Rather, following an investigation of human nature, Hobbes came to the conclusion that regardless of the diversity of passions among individuals, every sane person will have reason to fear death. Consequently, every person who would consult their reason would recognise the necessity of adhering to the laws of nature that provide the means towards self-preservation.

Finally, supplementing the textual evidence that presented above, we may provide an additional argument why our interpretation provides a more convincing reading of Hobbes's moral epistemology. As we have seen, the most common interpretation of Hobbes's account of practical deliberation concludes that humans are motivated by mindless and diverging passions, while the role of reason is restricted to finding the means to these desires. This interpretation therefore ascribes to Hobbes a rather pessimistic view of human nature, as these diverging passions will inevitably lead to conflicts that can only be resolved by subjecting to an all-powerful sovereign. Yet according to our own reading, Hobbes does not suppose that the ends of human beings are simply predetermined by passions following from their material constitution. Rather, Hobbes maintains that humans may employ their ability to reason about consequences in order to determine which ends are worth pursuing. Our reading thereby ascribes to Hobbes a somewhat less pessimistic and arguably more realistic view of human nature, which recognises that our actions are not simply determined by passions, but that humans are able to provide reasons for the ends they choose to pursue.

2. In Search for New Foundations: The Moral Epistemology of John Locke's *Essay*

Introduction

Like Thomas Hobbes before him, John Locke claims that human beings may perceive morality including natural law by means of reasoning. In the *Two Treatises of Government*, Locke in fact shows so much confidence in the power of reason that he thinks that even in the state of nature, human beings would largely abide by the laws of nature.[149] Yet in the *Two Treatises* themselves, Locke does not explain the precise method by which human beings may employ their reason to discern natural law. For an answer to this question, we must turn to Locke's *Essay concerning Human Understanding*, which is often primarily read as a treatise on epistemology, but also contains extensive discussions that touch upon ethical questions.[150] In fact, Locke himself states that he was motivated to write the *Essay* following discussions among friends about morality and revealed religion.[151] As the primary statement of Locke's theory of epistemology, the *Essay* thus not only explains how humans may acquire scientific knowledge, but also aims to outline a method for moral judgment. In this chapter, we will therefore study the *Essay* in order to reconstruct how Locke thinks that human beings should use their reason to grasp the universal morality of natural law.

Locke's approach to epistemology assumes that the question how to acquire true knowledge may be answered by revealing the mechanism by which

---

[149] Locke, *Two Treatises of Government* ed. by Ian Shapiro (New Haven 2003) II.ii.6 p.102.

[150] This chapter will use the *Essay* as its main source and only refer to other texts by Locke to provide context or clarification. We will therefore largely ignore the *Essays on the Law of Nature* which Locke wrote while a student at Oxford. Locke declined to publish these essays during his lifetime, which is not surprising in light of several major discrepancies with his mature writings. In addition, this chapter will also refer sparingly to the *Two Treatises of Government*, simply because these mainly deal with concrete questions of political right rather than moral epistemology. Finally, we will only briefly mention Locke's views on moral epistemology as expressed in the *Reasonableness of Christianity*, as he had clearly amended his position by the time he wrote this treatise.

[151] In the *Epistle to the Reader* of the *Essay Concerning Human Understanding*, John Locke recounts that he was instigated to write the treatise after a private discussion among friends. Although Locke himself does not specify the topic of this discussion, we know from one of the other participants that it was concerned with morality and revealed religion. Cf. Rogers, 'The Intellectual Setting and the Aims of the *Essay*' 8.

human beings collect knowledge in the first place.[152] The first step of Locke's methodology is therefore to uncover the ways in which human beings acquire ideas – the building blocks of their thoughts – and to trace the various means by which the mind may relate these ideas to produce knowledge. This account of human understanding would then also imply the correct method by which reason would be able to produce true knowledge. This approach provides Locke with a reasonably convincing account of the method by which human beings should acquire knowledge of empirical reality. Locke seeks to employ the same approach in order to explain how humans may grasp morality and natural law by means of reasoning. Thus Locke first describes how the mind acquires moral concepts, subsequently outlining the process by which reasoning with these concepts may produce normative conclusions.

Yet we will discover that Locke's theory of normative reasoning lacks the credibility of his account of the origin of empirical knowledge. Not only will we encounter several objections voiced by contemporaries, including Berkeley and Hume, against his account of moral reasoning. Modern scholars have also found it challenging to construct a coherent interpretation of Locke's moral epistemology. Nonetheless, in our subsequent discussion we will encounter a number of readings that have attempted to reconcile the paradoxes of Locke's moral philosophy. Yet our own investigation will proceed from the assumption that a valuable interpretation does not necessarily require an entirely coherent reconstruction of Locke's moral epistemology.[153] The interpretation proposed in this chapter will thus reconstruct Locke's account of moral epistemology as much as is warranted by textual evidence, but will not attempt to tie up all loose ends. Instead of an entirely coherent reconstruction, our interpretation will thus rather present a diagnosis of some of the inconsistencies and fissures of Locke's account of normative reasoning.

As Locke's account of moral reasoning is an integral part of his broader theory of human understanding, we will begin our discussion with a general

---

[152] Locke, *An Essay concerning Human Understanding* ed. by Peter Nidditch (Oxford 1975) I.i.2 p.43-44: 'It shall suffice to my present Purpose, to consider the discerning Faculties of a Man, as they are employ'd about the Objects, which they have to do with: and I shall imagine I have not wholly misimploy'd my self in the Thoughts I shall have on this Occasion, if, in this Historical, plain Method, I can give any Account of the Ways, whereby our Understandings come to attain those Notions of Things we have, and can set down any Measures of the Certainty of our Knowledge.'

[153] Schneewind, *The Invention of Autonomy* 142 proposes a similar approach to the interpretation of Locke's moral philosophy.

overview of his epistemology. This examination of Locke's theories on the origin of ideas and the properties of reason may appear more detailed than strictly necessary for our present concerns. Yet this overview will also serve as a point of reference for our discussions in later chapters, as we will see that Condillac, Helvétius, and Rousseau were all building upon Locke's theory of human understanding. After this overview, we will proceed with our investigation into Locke's moral epistemology by examining Locke's views on the origin of the moral concepts that comprise normative reasoning. Yet as we will find that Locke's account is hardly straightforward, we will discuss three scholarly interpretations that attempt a coherent reconstruction. The first of these ascribes to Locke the view that the composition of moral notions is derived from the conventional signification of moral terminology. A second line of interpretation claims that Locke assumed that human beings are endowed with an innate faculty of normative reason which allows them to grasp natural law. Finally, a third reading presents Locke as a consequentialist, who would maintain that natural law is nothing but a prudential rule that indicates the most effective means towards the satisfaction of the passions. While our own interpretation will approximate this consequentialist reading, we will conclude that it is not feasible to provide a coherent reconstruction of Locke's moral epistemology that does justice to all of his philosophical commitments and preconceptions.

The Origin of Ideas

John Locke's philosophy is predicated upon the fundamental thesis that all ideas in the mind can ultimately be traced back to experience – either in the form of sensation or reflection. In the previous chapter, we have seen that Hobbes had likewise accepted this premise. Yet whereas the position plays a relatively unassuming role in Hobbes's thought, Locke makes the sensory origin of ideas one of the central messages of his philosophy. Locke's emphasis on this position is mainly caused by his preoccupation with repudiating the theory of innate ideas – the notion, suggested by many Cartesian and some Scholastic philosophers that certain ideas are naturally present in the mind and do not need to be acquired.[154]

[154] It is beyond our present purposes to explore the concept of innate ideas in any detail. John Yolton, *Locke and the Way of Ideas* (Bristol 1993 [1957]) 30-47; Samuel Rickless, 'Locke's Polemic against Nativism', in: Lex Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* (Cambridge 2007) 33-66, 35-43 both provide an overview of the targets of Locke's argument.

At least according to Locke's interpretation, many of these philosophers held that such fundamental ideas as the notion of God or the basic rules of morality could or need not be acquired from experience, but were present – or at least potentially present – in the human mind at birth, having been placed there by God.[155] In the first book of the *Essay*, Locke is then mainly concerned with developing various arguments against the notion of innate ideas.[156]

While the details of these arguments need not concern us here, it is important to note that Locke also opposed the theory of innate ideas for moral reasons. The idea that morality had been inscribed on the mind by God himself, made it possible to claim that some morals were simply natural to man – and therefore self-evident and beyond question. This theory could then be employed to claim that certain moral conventions prevalent in society are beyond rational scrutiny because they are innate. Locke – who was an opponent of this kind of dogmatism – strongly dismissed this claim, stating that '*there cannot any one moral Rule be propos'd whereof a Man may not justly demand a Reason*: which would be perfectly ridiculous and absurd, if they were innate, or so much as self-evident.'[157] Locke thereby posited that no moral conventions or precepts should be exempt from rational scrutiny.[158] At the same time, Locke thereby dismisses the notion that moral judgment is some kind of intuitive perception of self-evident truths.[159] Rather, Locke states that 'moral Principles require Reasoning and Discourse, and some Exercise of the Mind, to discover the certainty of their Truth.'[160]

As Locke denies that any of the ideas we employ in reasoning – including normative reasoning – are either innate or self-evident, he needs to provide an alternative account of how human beings acquire or construct both their ideas of empirical objects and moral concepts. At the foundation of this account is Locke's famous axiom that all ideas of the understanding derive from either sense-

[155] Locke, *Essay* I.ii.1 p. 48.

[156] For a critical discussion of Locke's arguments, cf. Rafaella de Rosa, 'Locke's *Essay*, Book 1: The question-begging status of the anti-nativist arguments' *Philosophy and Phenomenological Research* 59 (2004) 37-64; Rickless, 'Locke's Polemic against Nativism' 45-66.

[157] Locke, *Essay* I.iii.4 p.68.

[158] This aspect of Locke's thought is particularly emphasised by Morton White, *The Philosophy of the American Revolution* (Oxford 1978) 9-23.

[159] Cf. Locke, *Essay* I.ii.8-9 p.51-52 for a more extensive rejection of reason as some kind of innate or intuitive faculty.

[160] Locke, *Essay* I.iii.1 p.66.

perception or reflection. Thus through sense-perception, the understanding will acquire ideas representing the objects of empirical reality, while reflection furnishes us with ideas of the operations of the mind. It is important to note that when Locke uses the term 'reflection,' he does not refer to what is commonly understood as reflection – some form introspective reasoning. Instead, for Locke 'reflection' is merely the passive perception of operation of the mind itself, which provides us with our ideas about operations like willing, judging or reasoning.[161] Through sense-perception and reflection, the mind is supplied with so-called simple ideas – singular impressions from individual objects or phenomena.[162] Examples include the ideas of colour, sound or taste, which are simple in the sense that they are the most fundamental units of thought, cannot be grasped by any linguistic definition, but solely acquired through experience.[163]

These simple ideas may subsequently be combined by the understanding to form several types of complex ideas that transcend the level of immediate experience and enable more advanced operations of the mind. According to Locke, all thought and cognition operates through ideas and concepts constructed from the simple sense impressions derived from the experience of either empirical reality or perception of the operations of the mind:

> Even *the most abstruse Ideas*, how remote soever they may seem from Sense, or from any operation of our own Minds, are yet only such, as the Understanding frames to it self, by repeating and joining together Ideas, that it had either from Objects of Sense, or from its own operations about them: So that even those large *and abstract* Ideas *are derived from Sensation, or Reflection,* being no other than what the Mind, by ordinary use of its own Faculties, employed about *Ideas*, received from Objects of Sense, or from the Operations it observes in it self about them, may, and does attain unto.[164]

Locke's insistence 'that even those large *and abstract* Ideas *are derived from Sensation'* has sometimes been taken out of context and interpreted as claiming that like simple ideas, complex ideas are simply acquired wholesale through sensation. This interpretation has been supported by Locke's somewhat

[161] Locke, *Essay* II.vi.1-2 p.127-128.

[162] Locke, *Essay* II.ii.1-3 p.119-121. Locke describes simple ideas as ideas that 'enter by the Senses simple and unmixed.' A distinctive feature of simple ideas is that they cannot be defined. Cf. Locke, *Essay* III.iv.7 p.422.

[163] Locke, *Essay* III.iv.4 p.421.

[164] Locke, *Essay* II.xii.8 p.166. Cf. II.i.2 p.104.

unfortunate adoption of the metaphor describing the mind as a *tabula rasa* – a blank slate. According to this reductionist interpretation, Locke sees the mind as a purely passive and receptive faculty that receives both its simple and complex ideas through sensation. Accordingly, the only real activity of the mind would be to reason with, and reflect on, the ideas it has already acquired through sensation.[165] Among those who interpreted Locke in this reductionist manner were some of the French *philosophes* of the eighteenth century. As we shall see in our discussion of Helvétius, their particular interpretation or adaptation of Locke's epistemology had important consequences for their conceptions of human nature.

Yet this interpretation fundamentally misunderstands Locke's theory. Thus Locke maintains that the understanding does not simply receive complex ideas, but needs to actively construct them. To be sure, in the subsequent section on the role of language in cognition, we will see that Locke envisions that complex ideas – and particularly moral concepts – may be learned from others through the medium of speech.[166] But at the same time, Locke maintains that the complex ideas that constitute our reasoning may also be constructed autonomously. The precise method of this construction varies for different types of ideas. Locke makes a strict distinction between complex ideas of substances, ideas of modes, and of relations. The first category contains ideas representing empirical objects acquired through sensation and/or reflection, such as a chair, tree or man. The second type of ideas of modes is not derived immediately from experience, but rather constructed by the mind at will from the simple ideas acquired through sensation or reflection. Ideas of modes allow for reasoning in two realms of knowledge not dependent on empirical investigation but nonetheless relevant to our daily lives: morality and mathematics. The ideas of the final category of relations indicate connections between singular complex ideas of either modes or substances. As this last category of ideas is not relevant to our purposes, we will limit our present discussion to complex ideas of substances and of modes, which we will now investigate in turn.[167]

---

[165] Nicholas Jolley, *Locke: his Philosophical Thought* (Oxford 1999) 28 notes that this is a common misconception.

[166] Locke, *Essay* III.iv.12 p.425-426.

[167] Cf. Martha Brandt Bolton, 'The Taxonomy of Ideas in Locke's Essay' in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 67-100, 92-94 for a more extensive discussion of ideas of relation. Relations of cause and effect are the most prominent example of complex ideas of relation between ideas of substances. But complex ideas of morality

As we have seen, complex ideas represent objects of empirical reality. They are constructed when the mind combines simple ideas customarily received in conjunction into a complex idea of an object:

> We come to have the Ideas of particular sorts of Substances, by collecting such Combinations of simple Ideas, as are by Experience and Observation of Men's Senses taken notice of to exist together, and are therefore supposed to flow from the particular internal Constitution, or unknown Essence of that Substance.[168]

Thus when we think or speak of a 'horse,' 'table' or 'cloud,' we evoke a complex idea that unites the simple ideas of sight, touch, smell, hearing, etc. previously acquired conjointly in experience into a single mental construct. For instance, the complex idea of a 'horse' is constructed by the understanding out the of simple ideas of the colour, shape, solidity, smell and sound previously experienced in objects classified as 'horses.' That the mind is not merely passive in the construction of these kinds of complex ideas, but in fact is actively involved in their assembly, is illustrated by the fact that the understanding even has the capacity to combine simple ideas into new and even fantastical combinations. This explains our ability to conjure up both new inventions as well as fictional beings such as unicorns.[169]

As he claims that complex ideas are not simply received through sensation but rather actively constructed, Locke implies the existence of an autonomous faculty of the mind that operates independent from the immediate stimuli of sensation. Indeed, throughout the *Essay*, we find mentions of the ability of the mind to actively inspect, relate, compose, and decompose its ideas. Yet while Locke describes at length the origin and properties of ideas, he declines to reveal the source or characteristics of the autonomous faculty that seems to have such a crucial role in his account of human understanding. Rather, it appears that Locke simply assumes – albeit tacitly – the presence of this autonomous perceiving and assenting faculty of the mind. Accordingly, modern critics of empiricism including Richard Rorty have pointed out that as Locke's account of human understanding tacitly relies on this 'ghostly entity', his

---

may also exhibit mutual relations. Thus the complex idea of 'father' has a necessary relation to the idea of either 'son' or 'daughter.' Locke envisages that the perception of these relations between complex ideas of either modes or substances produce this third type of complex ideas of relations.

[168] Locke, *Essay* II.xxiii.3 p.296.

[169] Locke, *Essay* II.xxiii.1-14 p. 295-305.

epistemology ultimately eludes its goal of describing how the mind may transform sensations into knowledge representative of empirical reality.[170] Yet for our present purposes it is more important to note that Locke's ambiguity on the issue was also recognised by some of his early-modern readers, including Condillac. Thus in our next chapter we will see that Condillac sought to remedy the deficiency in Locke's theory by devising an empiricist account of human understanding that no longer required any autonomous faculty of the mind. Conversely, in our final chapter, we will see that Rousseau adopts many elements of Locke's theory of human understanding, but not without making explicit the premise that human reason relies on an autonomous faculty of judgment – albeit without further explaining the origin or properties of this faculty.

Despite the apparent active involvement of the mind in the composition of its ideas, Locke's credo that all ideas in the mind come from sensation or reflection entails that our comprehension and acquaintance of objects of empirical reality will never go beyond what we are able to sense about them. As is also evident from the passage quoted above, Locke assumes that the sensible qualities we perceive in an object are the consequence of an 'unknown essence.' Inspired by his friend Robert Boyle, Locke envisions that physical objects consist of corpuscles – or atoms, as they would nowadays be called – of which the properties and organisation are responsible for the sensible properties of the objects they comprise.[171] But the actual nature of this corpuscularian essence – usually referred to as an object's 'real essence' – will necessarily remain hidden from us, as it will only reveal itself indirectly through sensible qualities.[172] This position that the real essences of substances are beyond our comprehension is generally referred to as Locke's doctrine of 'essence agnosticism.'

It is important to note that Locke's essence agnosticism diverges distinctly from the position of many of his contemporaries or predecessors – Scholastics and Cartesians alike. Many of these thinkers maintained that humans

---

[170] Rorty, *Philosophy and the Mirror of Nature* 144: 'since for him [i.e. Locke] impressions were *representations*, he needed a faculty which was *aware* of the representations, a faculty which *judged* the representations rather than merely *had* them – judged that they existed, or that they were reliable, or that they had such-and-such relations to other representations. But he had no room for one, for to postulate such a faculty would have intruded a ghost into the quasi-machine whose operation he hoped to describe.'

[171] Cf. Steven Forde, *Locke, Science and Politics* (Cambridge 2014) on the influence of Boyle on Locke.

[172] Locke, *Essay* II.viii.7 p.134; II.xxiii.29 p. 312; IV.iii.11-14 p. 544-546.

have the ability to grasp the real essences of substances, if only they use the correct methods of investigation.[173] Thus Locke's contemporary critic John Sergeant maintained that 'The impressions from objects that affect the senses, and by them the soul, do carry the very nature of those objects along with them. As a consequence, individuals have the ability to mould their ideas and concepts to represent the *species* and *genera* present in nature.'[174] As these thinkers assume that human beings are able to understand the true nature of the objects of empirical reality, they envision that we can also acquire demonstrative and certain knowledge of these objects.[175]

Conversely, Locke's doctrine of essence agnosticism rules out this possibility. As we are unable to ascertain the real essences of objects, our complex ideas of substances solely unite the sensible properties of objects into what Locke refers to as a 'nominal essence.'[176] But because sensory perception is inherently subjective, individuals are prone to frame more or less diverging nominal essences of objects. Of course, humans have a similar physical constitution and possess the same sensory organs, and they will not perceive objects in radically different ways. The nominal essences of ideas of substances will thus tend to display at least superficial correspondence among individuals.[177] Nonetheless, Locke thinks that there is still considerable latitude for variation:

> But however, these *Species* of Substances pass well enough in ordinary Conversation, it is plain, that this complex *Idea*, wherein they observe several Individuals to agree, is, by different Men, made very differently; by some more, and others less accurately.[178]

---

[173] There is reason to believe that despite, the enormous advances in the natural sciences of the past centuries Locke's principle of essence agnosticism is still not refuted by modern science. Even now that the inner structure of atoms has been uncovered, the boundaries of our knowledge of particle physics are continuously pushed further, but with each discovery new questions arise. The fundamental nature of matter thus still remains a mystery. Cf. Margaret Atherton 'Locke on Essences and Classification' in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 258-285, 267-269.

[174] John Sergeant, *The Method to Science* (London 1696) 2; Cf. J.W. Yolton, 'Locke's Unpublished Marginal Replies to John Sergeant', *Journal of the History of Ideas* 12 (1951) 528-559; Dmitri Levitin, 'Reconsidering John Sergeant's Attacks on Locke's *Essay*,' *Intellectual History Review* 20 (2010) 457-477.

[175] Cf. Michael Ayers, *Locke I: Epistemology* (London and New York 1991) 26-35.

[176] For a discussion of Locke's view on *species* and *genera*, see Atherton 'Locke on Essences and Classification.'

[177] Locke, *Essay* III.vi.28-29 p.455-456.

[178] Locke, *Essay* III.vi.31p.458.

Consciously or unconsciously, people will thus construct their nominal essences of objects in more or less diverging ways. Through observation or empirical investigation nominal essences may be refined, but they will never be determined definitively by being made correspondent with real essences – simply because the latter are inaccessible.[179]

Locke's position on the elusiveness of the real essences of substances also limits our ability to gain certain knowledge of empirical reality. Our limitation to the nominal essences of substances means that we may be able to document the action and operation of substances on the basis of our experiences of them, but we will never be able to truly understand the mechanism responsible for these experiences. Because of this inability to grasp the inner functioning of nature, Locke thinks that certain and indisputable knowledge of empirical reality will necessarily remain beyond our grasp. Reasoning with ideas of substances will therefore not proceed beyond the status of probability. This does not mean that investigation of empirical reality is fruitless – the resulting probabilistic knowledge generally proves to be very useful and may be regarded as true until proven otherwise.[180] Locke merely wishes to emphasise that because of our limitation to sensible qualities and ignorance of real essences, our knowledge of substances simply cannot attain the standard of absolute certainty that Sergeant and many scholastics presumed possible.

Locke's essence agnosticism and resulting conclusion that all knowledge of empirical reality is probable at most may appear to lead to a form of distressing scepticism. But Locke is careful to provide a theological explanation for the limitation of man's cognitive abilities. Thus while human beings cannot gain demonstrative knowledge of empirical objects, this is not at all problematic, as God has ensured that man's cognitive abilities are not only sufficient, but even calibrated to the requirements of life on earth.[181] Locke therefore claims that the primary purpose of the understanding is not the full comprehension of the natural world, but rather to know how humans should govern their actions:

> Our Business here is not to know all things, but those which concern our Conduct. If we can find out those Measures, whereby a rational Creature put in that State, which

[179] Locke, *Essay* II.xxiii.3-5 p.296-298; IV.iii.11-17 p. 544-548. Cf. Jolley, *Locke* 150-155.
[180] Locke, *Essay* IV.iv.12-18 p. 568-573.
[181] Locke, *Essay* I.i.5 p.45

> Man is in, in this World, may, and ought to govern his Opinions, and Actions depending thereon, we need not be troubled, that some other things escape our Knowledge.[182]

In fact, Locke maintains that a neglect of man's intellectual limitations is prone to induce unrestrained investigations, producing nothing but trivial disputes and unfounded theories, ultimately leading to scepticism rather than knowledge.[183] Although this does not entail that we should ignore the pursuit of the natural sciences, Locke thinks that we should be satisfied with probable knowledge of empirical reality.

At the same time, Locke suggests in the passage quoted above that God has equipped man's intellectual faculties with the ability to know how one should govern his actions. In other words, God has ensured that human beings are able to acquire knowledge of morality. And in contrast to empirical science, this knowledge of morality is capable of demonstrative certainty. Locke does not solely support this position by referring to Divine intentions. He also presents several additional reasons why human beings have the ability to gain demonstrative knowledge of morality. At the most fundamental level, his reason is that normative reasoning is conducted through so-called ideas of modes which do not depend on our experience of empirical reality. This type of ideas constitutes or enables two distinct fields of inquiry – mathematics and ethics. As the discussion of mathematics is beyond the scope of our current investigation, we will focus on the so-called mixed modes that comprise normative reasoning. Locke himself describes ideas of modes as:

> such complex *Ideas*, which however compounded, contain not in them the supposition of subsisting by themselves, but are considered as Dependences on, or Affections of Substances; such are the *Ideas* signified by the Words *Triangle*, *Gratitude*, *Murder*, etc.[184]

---

[182] Locke, *Essay* I.i.6 p.46.

[183] Locke, *Essay* I.i.7 p.47: 'Thus Men, extending their Enquiries beyond their Capacities, and letting their Thoughts wander into those depths, where they can find no sure Footing; 'tis no Wonder, that they raise Questions, and multiply Disputes, which never coming to any clear Resolution, are proper only to continue and increase their Doubts, and to confirm them at last in perfect Scepticism.'

[184] Locke, *Essay* II.xii.4 p.165.

Like complex ideas of substances, ideas of modes are constructed by the mind on the basis of simple ideas. But whereas ideas of substances are collections of simple ideas habitually associated with specific objects, modes are concepts freely created by the mind independent from experience.[185] Locke refers to moral ideas specifically as 'mixed-modes' – signifying the fact that while these concepts are archetypes freely constructed by the mind and do not refer to any concrete object, they are nonetheless assembled from ideas acquired through sensation or reflection and may be projected upon the objects of empirical reality.[186] Thus complex ideas such as 'property,' 'citizen,' 'father,' 'justice' or 'murder' do not have any essential existence in empirical reality. They are concepts designed by the mind and rather projected upon reality. Concepts like property and murder may be identified by us in the events and objects we encounter, but these concepts correspond to no particular empirical objects and have no real existence outside the realm of thought and language.

Moral Demonstration

Locke thus maintains that mixed modes are 'the Workmanship of the Mind' and may be constructed by individuals at will.[187] Yet he thinks that reasoning with these moral concepts may nonetheless yield useful and concrete knowledge about morality. This process of reasoning involves a comparison and relating of ideas to each other, whereby 'Reason perceives the necessary, and indubitable connexion of all the *Ideas* or Proofs one to another, in each step of any Demonstration that produces Knowledge.'[188] By diligently applying our reason to the consideration of mixed modes, Locke thinks that we may understand morality with demonstrative certainty:

[185] As Locke himself phrases it at II.xxxi.3 p. 376, modes are 'voluntary collections of Simple Ideas which the Mind puts together, without reference to any real archetypes or standing patterns existing anywhere.' Cf. III.v.3-6 p.429-431.

[186] An alternative definition is given by Steven Forde, '"Mixed Modes" in John Locke's Moral and Political Philosophy' *The Review of Politics* 73 (2011) 581-608, 587: 'A mode, then, to repeat, is an idea of an attribute that is separable from ideas of objects, but not conceived as having any existence independent of objects, ideas like *dozen*, or *rolling*.'

[187] Locke, *Essay* III.v.4 p.430. In fact, mixed modes are not the only type of ideas referred to as 'workmanship of the mind' by Locke. Thus he uses the same phrase to describe universal ideas of species, which are indeed also created by the mind at will.

[188] Locke, *Essay* IV.xvii.2 p.669.

> *Morality is capable of Demonstration*, as well as Mathematicks. For the *Ideas* that Ethicks are conversant about, being all real Essences, and such as, I imagine, have a discoverable connexion and agreement one with another; so far as we can find their Habitudes and Relations, so far we shall be possessed of certain, real, and general Truths: and I doubt not, but if a right method were taken, a great part of Morality might be made out with that clearness, that could leave, to a considering Man, no more reason to doubt, than he could have to doubt the Truth of Propositions in Mathematicks, which have been demonstrated to him.[189]

According to Locke, the reason why human beings may gain certain knowledge of morality is not merely that God intends us to do so. His main argument is that moral reasoning employs ideas of modes of which we may grasp the real essence. Unlike ideas of substances, moral and mathematical concepts are human constructions not tied to any external object with an unknowable real essence.[190] Accordingly, we are able to grasp the real essences of modes simply because – unlike with ideas of substances – real essence and nominal essence coincide.[191] Consequently, reason has the ability to perceive connections between moral concepts with complete accuracy, leading to knowledge of demonstrative certainty.

Following his theory of moral reasoning, Locke provides a few examples of how moral reasoning functions in practice. The first example of a moral demonstration Locke considers 'as certain as any demonstration in Euclid,' is the proposition that 'where there is no property, there is no injustice.'[192] Locke asserts that the concept of property means the tenancy of exclusive rights to a

[189] Locke, *Essay* IV.xii.8 p.643-644. Cf. IV.iii.18 p.549; IV.iv.7 p.565; I.iii.1p.66.

[190] Locke, *Essay* IV.iv.5 p. 564: '*All our complex* Ideas*, except those of Substances,* being *Archetypes* of the Mind's own making, not intended to be the Copies of any thing, nor referred to the existence of any thing, as to their Originals, *cannot want any conformity necessary to real Knowledge*. For that which is not designed to represent any thing but it self, can never be capable of a wrong representation, nor mislead us from the true apprehension of any thing, by its dislikeness to it.' Cf. IV.xii.6 p. 642.

[191] Locke, *Essay* III.iii.18 p.418. Jolley, *Locke* 155 correctly notes that this is the crucial aspect distinguishing complex ideas of modes from complex ideas of substances – and not the fact that the first type is constructed by the understanding and the second type supposedly not. As Jolley points out, both types of complex ideas are constructs of the understanding. See 155-157 for some potential objections to Locke's claim that real essences of ideas of modes can be known. Cf. Leibniz, *New Essays on Human Understanding* ed. by Peter Remnant and Jonathan Bennett (Cambridge 1996) III.iii.15-19 pp. 293-296.

[192] Locke, *Essay* IV.iii.18 p. 549.

certain thing, while injustice entails the infringements of these rights. Consequently, when property rights are abolished or absent, they can no longer be infringed, making injustice inconceivable. Locke's second example works in a similar way when he states that 'no government allows absolute liberty.'[193] Here the reasoning is that the restriction of liberties through laws is inherent in the concept of government. Any government that would allow absolute liberty would therefore be inconceivable. Locke suggests that through the diligent application and further extension of these kinds of demonstrations one could construct a veritable system of moral science.[194]

Yet it is not certain that this claim stands up to scrutiny. Both contemporary and modern commentators have reacted with scepticism to Locke's claim for the possibility of a moral science based on demonstrations like the examples he provides – and for several good reasons.[195] First of all, it is not entirely certain whether the two deductions Locke provides as examples actually produce any new knowledge. Thus when Locke maintains that there is no injustice without property, he is merely stating a conclusion that is already contained within the concepts themselves. His deductions therefore tell us something about the terms involved, but do not produce any further knowledge. In Kantian terms, Locke's deductions may be seen as analytic rather than synthetic. This becomes even more conspicuous if we restate Locke's deductions as propositions. Thus the first example could be rephrased as 'injustice is the infringement of property,' while the second deduction could simply be read as 'government restricts liberty.' Following David Hume, modern commentators have noted that these reformulations suggest that Locke's supposed demonstrations are in fact little more than explications of terms.[196] Because they

[193] Locke, *Essay* IV.iii.18 p. 550.

[194] Locke, *Essay* IV.iii.18 p. 549, where Locke suggests that in light of our idea of God and ourselves as rational beings, we 'might place *Morality amongst the Sciences capable of Demonstration*: wherein I doubt not, but from self-evident Propositions, by necessary Consequences, as incontestable as those in Mathematicks, the measures of right and wrong might be made out.'; Cf. IV.iv.5-9 p.564-567.

[195] Cf. J.L. Mackie, *Problems from Locke* (Oxford 1976) 208-209.

[196] David Perry, 'Locke on Mixed Modes, Relations, and Knowledge', *Journal of the History of Philosophy* 5 (1967) 219-235, 234-235 maintains that in Locke's own terminology, these demonstrations qualify as 'trifling propositions.' Hume, *An Enquiry concerning Human Understanding* XII p.163 recognises the same point, albeit without mentioning Locke by name, when he states: 'To convince us of this proposition, *that where there is no property, there can be no injustice*, it is only necessary to define the terms, and explain injustice to be a violation of property.

do not yield any knowledge beyond that contained in the terms employed, it is hard to imagine how demonstrations like those of Locke's examples could lead to a concrete science of morals.[197] This judgment was shared by Locke's contemporary George Berkeley, who in reaction to Locke's ethics quipped that to 'demonstrate Morality it seems one need only make a Dictionary of Words and see which included which. This is the greatest part and bulk of the work.'[198]

As Locke's theory indicates that moral reasoning is conducted by comparing and connecting mixed modes, this means that the outcome of demonstrations is determined by the composition of the moral concepts involved. Consequently, it seems that the fundamental and decisive part of normative judgment is not demonstration, but rather the composition of the moral concepts it involves. At first glance, Locke's account of moral demonstration suggests that he might have underestimated the fact that the composition of moral concepts is neither self-evident nor uncontested.[199] In order to provide a complete theory of moral reasoning, Locke would thus have to specify the method by which human beings may accurately compose their mixed modes. We have already seen that as modes, moral concepts are not derived from any objects of experience. Rather they may be constructed by the mind at will. This suggests that persons are free to construct their mixed modes in any way they please, leading to diverging results in moral reasoning. Yet we have seen that Locke considers morality capable of demonstration – suggesting that morality is dependent on a standard independent from individual perspectives or preferences. In order to understand how Locke considers morality capable of demonstration, we need investigate in greater detail the method by which Locke thinks individuals acquire or compose the moral concepts they employ in their reasoning. In the following, we will then first see that in Locke's theory, moral concepts are dependent on language, as

---

This proposition is, indeed, nothing but a more imperfect definition.' Cf. Hume, *Enquiry concerning the Principles of Morals* ed. by Tom L. Beauchamp (Oxford 1998) appendix I.ii §11 p.86.

[197] Leibniz also hints at this point when he states in the *New Essays* IV.iii p.385 that Locke's second demonstration 'belongs among the corollaries, i.e. the propositions which have only to be brought to one's attention [for their truth to be recognized].'

[198] Quoted in: Catherine Wilson, 'The Moral Epistemology of Locke's *Essay*', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 381-405, 397. Cf. Schneewind, *The Invention of Autonomy* 151n.

[199] Wilson, 'The Moral Epistemology of Locke's *Essay*', 397 also makes this point. Yet in *Essay* III.ix.7 p.478 Locke does display some awareness of the contested nature of moral notions, although he seems to be mainly concerned with the contested nature of moral *terms*.

mixed modes cannot exist without the use of corresponding moral terms with a certain conventional signification. Subsequently, we will turn to the question how moral concepts may be composed independent from the conventions of language in order to produce the universal morality of natural law.

The Signification of Moral Terms

Thus far, our discussion has focused on Locke's classification of ideas in the minds of individuals and the particular roles of ideas in various types of reasoning. Yet while Locke never tires to emphasise that the contents of the mind may always be traced back to experience, he recognises – although belatedly, by his own admission – that language also plays an important role in the way human beings frame their ideas.[200] According to Locke, language then has three main functions. On the one hand, human beings use words with a definition circumscribed by conventions of common use to communicate their ideas to each other. Yet on the other hand, Locke thinks that words also have an important role in cognition as structural aids in the composition of ideas – in particular in the case of mixed modes. Locke therefore suggests that moral concepts and moral terms mutually depend upon each other. Thirdly, Locke thinks that by the use of words, it is possible to acquire complex ideas from others, independent from our own experience.[201] Consequently, there are some grounds to ascribe Locke the view that individuals often derive the composition of their mixed modes from the conventional signification of moral terms. But before we may turn to this issue of the relationship between language and morality, we must first understand some of the most important premises of Locke's philosophy of language. We will therefore begin by reviewing Locke's theory of signification, and subsequently consider how linguistic conventions enable communication. Only then will we be in the position to determine the extent to which the composition of moral concepts is dependent upon the words to denote them in common conversation.

[200] Locke, *Essay* III.ix.21 p.488 for this admission.

[201] Locke, *Essay* III.iv.12 p.425-426: 'The case is quite otherwise *in complex Ideas*; which consisting of several simple ones, it is in the power of Words, standing for the several *Ideas*, that make that Composition, to imprint complex *Ideas* in the Mind, which were never there before, and so make their Names be understood. In such Collections of *Ideas*, passing under one name, *Definitions*, or the teaching the significations of one word, by several others, has place, and *may make us understand the Names of* Things, which never came within the reach of our Senses; and frame *Ideas* suitable to those in other Men's Minds, when they use those Names.'

The fundamental – and arguably the most distinctive – claim of Locke's philosophy of language is that '*Words in their primary or immediate Signification, stand for nothing, but the* Ideas *in the Mind of him that uses them*.'[202] Thus when a language user utters or comprehends a word, the meaning of this word is determined by a corresponding idea in his or her mind. According to Locke this means that words do not immediately signify the empirical objects we encounter in experience. He maintains that words do not stand for things, but only for the ideas of things that we have framed in our mind. This position follows partly from Locke's essence agnosticism – the view that the real essences of substances will always remain beyond human comprehension. Locke then argues that words cannot possibly signify anything of which we cannot have an idea:

> Till he has some *Ideas* of his own, he cannot suppose them to correspond with the Conceptions of another Man; nor can he use any Signs for them: For thus they would be the Signs of he knows not what, which is in Truth to be the Signs of nothing. But when he represents to himself other Men's *Ideas*, by some of his own, is he consent to give them the same Names, that other Men do, 'tis still to his own *Ideas*; to *Ideas* that he has, and not to *Ideas* that he has not.[203]

Thus when we speak about an empirical object, we only refer to the nominal essence the mind has constructed from our own personal experience of the object. This nominal essence contained in the idea is of course itself a representation of an object. Therefore we may ascribe Locke the view that while words *immediately* signify ideas, they *mediately* signify objects of which the immediately signified ideas are representations.[204]

---

[202] Locke, *Essay* III.ii.2 p.405. Cf. III.i.1 p.402.

[203] Locke, *Essay* III.ii.2 p.406.

[204] Michael Losonsky, 'Language, Meaning, and Mind in Locke's *Essay*', in: Lex Newman (ed.), *The Cambridge Companion to Locke's Essay* (Cambridge 2007) 286-312, 289 maintains that this is the most common interpretation among modern scholars, first developed in Norman Kretzmann, 'The Main Thesis of Locke's Semantic Theory', *The Philosophical Review* 77 (1968) 175-196. Jennifer Ashworth, '"Do Words Signify Ideas or Things?" The Scholastic Sources of Locke's Theory of Language', *Journal of the History of Philosophy* 19 (1981) 299-326 provides an alternative interpretation of Locke's theory of signification by placing it in the context of scholastic discussions of language. She suggests that in scholastic disputations – of which some were probably familiar to Locke – the term *significare* may indicate equal relations of signification between terms on the one hand and both ideas in the mind and things on the other. For a critique of this interpretation, see

The precise consequences of Locke's account of signification emerge only when we contrast it with other contemporary theories. Locke diverges most notably from the realist position of Aristotelianism. Scholastic and other realist early-modern philosophers claimed that human beings do have the ability to ascertain the real essences of objects. Accordingly, realist philosophers held that particular terms signify both an idea in the mind as well as the real essence of the empirical object from which the idea is an exact representation. In addition, these realist philosophers generally maintained that natural objects themselves are ordered into distinct *species* and *genera*. By perceiving these categories of objects, human beings have the ability to order their ideas in correspondence with the classes of objects that exist in nature. Consequently, realist philosophers also thought that universal terms referred both to the *species* and *genera* present in nature, as well as to ideas of these forms present in the mind.[205] Following this theory of signification, realist thinkers held that in theory, language could function as a flawless method of communication, when the real essences of objects and relations of signification would be illuminated and settled in the minds of all.[206]

As Locke renounces the notion that nature is ordered into clearly distinguishable *species* and *genera*, and rather claims the real essences of objects are inaccessible to human understanding, he considers this kind of realism as unfounded.[207] Consequently, he maintains that it is misguided to presume that words stand for anything but ideas in our own mind.[208] Yet Locke's own theory of signification has not been without its critics. Their main objection is that while Locke's theory describes how the understanding assigns meaning to words, it has difficulty explaining how language can also be used successfully in communication. The main issue is that because words immediately signify solely ideas in the minds of individuals, it is hard to see how they could serve as a

---

Michael Losonsky, 'Locke on Meaning and Signification', in: G.A.J. Rogers (ed.), *Locke's Philosophy, Content and Context* (Oxford 1994) 123-142.

[205] For a more extensive discussion of Locke's views on *species* and *genera* in the context of Aristotelian and other early-modern theories, Cf. Michael Ayers, *Locke II: Ontology* (London and New York 1991) 65-77.

[206] Some realist thinkers thus presumed that language could be reformed to become a perfect representation of reality. Cf. Dawson, *Locke, Language and Early-Modern Philosophy* 102-107.

[207] Locke, *Essay* III.iii.12 p.415.

[208] Locke, *Essay* III.ii.5 p.407: 'Though give me leave here to say, that it is a perverting the use of Words, and brings unavoidable Obscurity and Confusion into their Signification, whenever we make them stand for any thing, but those *Ideas* we have in our own Minds.'

means to transfer our precise thoughts to each other. Thus according to Locke's theory, when a speaker utters a term, a listener will assign meaning to this word by means of an idea present in her own mind, rather than an idea in the mind of the speaker. Because the ideas language users attach to a term need not necessarily correspond, they will often attach diverging significations to the words they use in communication. By itself, Locke's theory of signification therefore suggests that individuals all have different views on the meaning of words, and thereby effectively speak a private language that would preclude the possibility of accurate communication.[209]

This conclusion is particularly disconcerting in the case of moral terms. We have seen that following his essence agnosticism, Locke thinks that individuals often construct diverging ideas of empirical objects. Nonetheless, the correspondence of our sense-organs and similarity in our experience of empirical reality causes at least a superficial measure of conformity in ideas of substances among individuals. Consequently, the signification of terms denoting empirical objects would show only a limited extent of divergence among individuals.[210] Yet as 'assemblages of *Ideas* put together at the pleasure of the mind,' ideas of modes are not based on any models in nature, and would therefore appear to lack an independent standard by which they could be constructed. Locke therefore maintains that the names of mixed modes are often 'very various and doubtful.'[211] The worry that individuals retain private notions of the meaning of words is therefore especially pertinent to moral language.

Common Use and the Meaning of Words

The signification of words denoting substances thus derives some superficial degree of uniformity from the fact that objects occasion similar sensations in all individuals. Yet Locke identifies a second way in which the signification of terms – both of substances and of modes – receives a degree of conformity sufficient to facilitate communication among individuals. Locke recognises that while terms immediately signify nothing but ideas in the understanding of individuals, the

[209] Hannah Dawson, 'Locke on Private Language', *British Journal for the History of Philosophy* 11 (2003) 609-638; Dawson, *Locke, Language and Early-Modern Philosophy* 239-276.

[210] Locke, *Essay* III.vi.28-29 p.455-456; Cf. Dawson, *Locke, Language, and Early-Modern Philosophy* 220.

[211] Locke, *Essay* III.ix.478.

meaning of words is also governed or at least circumscribed by linguistic conventions:

> For words, especially of Languages already framed, being *no Man's private possession*, but the common measure of Commerce and Communication, *'tis not for any one, at pleasure, to change the Stamp they are current in*; nor alter the Ideas they are affixed to.[212]

Thus in principle, individuals are free to construe the signification of words in a more or less idiosyncratic way. But in order to be understood, interlocutors will have to accommodate the usage of their terms at least roughly to some pre-existing and socially established conventional meanings. Due to this requirement, individuals are forced, and will often try, to reconcile the signification of their words to the standards of common use.

Apart from the uniformity of the senses and similarities in experience, the necessity to conform to certain linguistic conventions is then an important factor in generating conformity in the signification of terms among individuals. According to Martin Lenz, this role of common use in Locke's philosophy of language is in fact sufficient to nullify the objection that if words immediately signify ideas this entails that individuals each speak a private language. Lenz maintains that the standards of common use are stringent and precise enough to alleviate or even dissolve the most serious of the sceptical consequences drawn by some interpreters of Locke's theory of signification. While Lenz thereby removes an important objection to Locke's philosophy of language, he does so by interpreting it as a social externalist theory. This means that Lenz ascribes Locke the position that the meaning of words – and thereby the composition of our ideas – is determined to a significant extent by shared linguistic conventions. Lenz therefore ascribes to Locke the view that our thoughts are at least partly determined by the constitution of our language.[213]

Lenz is certainly correct to point out that conventions of common use play a crucial role in enabling communication by circumscribing the signification of words among language users. But it is not clear that an appeal to linguistic convention can entirely dissolve the objection that Locke's theory of signification leads to a conception of language as a private construction. In fact, Locke himself

[212] Locke, *Essay* III.xi.11 p.514.

[213] Martin Lenz, 'Locke as a Social Externalist', *Proceedings of the British Academy* 189 (2013) 53-67; Cf. Lenz, *Lockes Sprachkonzeption* (Berlin and New York 2010).

warns that linguistic conventions can only provide limited stability and accuracy to the signification of words:

> 'Tis true, *common Use*, that is the Rule of Propriety, may be supposed here to afford some aid, to settle the signification of Language; and it cannot be denied, but that in some measure it does. Common use *regulates the meaning of Words* pretty well for common Conversation; but no body having an Authority to establish the precise signification of Words, nor determine to what *Ideas* any one shall annex them, common Use is not sufficient to adjust them to philosophical Discourses.[214]

Thus while Locke thinks that common use is sufficient to facilitate everyday conversation, linguistic conventions lack the required precision and accuracy for philosophical discussion – which according to Locke's terminology would also include discussion on topics of empirical science. Locke's first reason for this position is that the conventions of common use have generally been established in times and by people ignorant of science and philosophy.[215] These flawed conventions of language are tenacious, as they are socially constructed and cannot be amended or changed by an individual without the consent of others.[216] The conventions of common use are therefore permeated with remnants of the ignorance and prejudices of earlier ages.[217] We will see below that Condillac concluded from this suggestion that scientific and moral progress required the wholesale reform of the conventions of language – a conclusion not drawn by Locke himself.

A second reason why common use is an insufficient regulator of the meaning of words is that its conventions are not only often inaccurate, but also 'doubtful and uncertain.' In particular in the case of very complex ideas, such as universal ideas of substances or moral concepts, the conventions of language are not sufficiently settled and circumscribed to ensure successful communication and preclude the possibility that speakers misunderstand each other because they

---

[214] Locke, *Essay* III. ix.8 p.479. Cf. III.ii.8 p.408; III.vii.51 p.471; III.ix.15 p.484.

[215] Locke, *Essay* III.vi.25 p.452-453: 'Since Languages, in all Countries, have been established long before Sciences. So that they have not been Philosophers, or Logicians, or such who have troubled themselves about *Forms* and *Essences*, that have made the general Names, that are in use amongst the several Nations of Men: But those, more or less comprehensive terms, have, for the most part, in all Languages, received their Birth and Signification, from ignorant and illiterate people, who sorted and denominated Things, by those sensible Qualities they found in them.'

[216] Locke, *Essay* III.vi.51 p.470-471.

[217] Cf. Dawson, *Locke* 267-269.

attach different ideas to the same terms.[218] Despite the role of linguistic conventions in regulating the signification of words among individuals in everyday conversation, the objection that speakers may use a private language therefore remains applicable to scientific and philosophical discussions. Ultimately, the standard of common use is thus insufficient to entirely negate the sceptical consequences of Locke's thesis that words immediately signify nothing but ideas in the mind. In fact, Locke himself seems to think that despite linguistic conventions, language nonetheless remains an imperfect method of communication – especially concerning morality.[219]

Words and Ideas

The social externalist reading proposed by Lenz is thus not entirely able to invalidate the private language objection levelled at Locke's theory of signification. But Lenz also claims that for Locke, the conventional signification of words determines or at least influences the composition of our ideas. There is at least one significant reason why at first glance, this reading might make sense. Like Hobbes, Locke thinks that language plays an important role in cognition as words are indispensible to the composition and classification of complex ideas. But while Locke agrees with Hobbes that there is a close relationship between words and complex ideas, he diverges from his predecessor in at least one significant respect. In the previous chapter, we have seen that Hobbes maintains that abstract and universal *ideas* are by themselves inconceivable, and that abstractions and universals only exist in the mind in the form of words with a certain reference specified by a definition.[220] By contrast, Locke does think that humans may construct abstract or universal ideas by compounding particular ideas and isolating them from their circumstances – a view that would later be

[218] Locke, *Essay* III. ix.8 p.479: 'From all which, it is evident, that the Name of such kind of very complex *Ideas*, are natural liable to this imperfection, to be doubtful and uncertain signification; and even in Men, that have a Mind to understand one another, do not always stand for the same *Idea* in Speaker and Hearer.' Cf. Locke, *Essay* III. ix.21 p.489; III.ix.4 p.476-477; III.ix.6 p.478.

[219] Locke, *Essay* III. xi.4 p.510: 'For he that shall well consider the *Errors* and Obscurity, the Mistakes and Confusion, that is *Spread in the World by an ill use of Words*, will find some reason to doubt, whether Language, as it has been employ'd, has contributed more to the improvement or hindrance of Knowledge amongst Mankind. How many are there, that when they would think on Things, fix their Thoughts only on Words, especially when they would apply their Minds to Moral Matters?' Cf. Dawson, *Locke* 234-235 who agrees with this estimation.

[220] Locke, *Essay* III.iii.12 p.415.

criticised as unintelligible by Berkeley.[221] In contrast to Hobbes, Locke therefore thinks universal and abstract terms signify a complex idea in the mind – an abstracted mental image. While Hobbes's theory can be described as a form of nominalism, Locke's position may rather be considered as conceptualism.[222]

Although Locke thinks that ideas and words are separate entities, he does recognise that in the case of abstract or general ideas, the former generally depend upon the latter.[223] Thus Locke thinks that words have an important cognitive function in structuring, delineating and preserving complex ideas:

> He that has complex *Ideas*, without particular names for them, would be in no better Case than a Bookseller, who had in his Ware-house Volumes, that lay there unbound, and without Titles.[224]

Although Locke does not spell it out, this passage suggests that words function as cognitive labels that aid the understanding in memorising and retrieving compounded ideas. According to Lenz, this means that for Locke, the mind cannot retain or process complex ideas of both substances and modes without the use of words. Upon this reading, words and ideas would therefore mutually depend upon each other, with words acquiring their signification from ideas, while ideas rely on words for their stability and accessibility to the understanding. Consequently, Lenz maintains that the composition of ideas is influenced or determined by the conventional signification of the words upon which these ideas depend.[225]

On the basis of the passage quoted above, Lenz' interpretation might seem appealing. But a closer inspection of Locke's *Essay* reveals that at least one aspect of Lenz' reading that discerns a mutual dependency of words and ideas

---

[221] Locke, *Essay* III.iii.6 p.411: '*Ideas* become general, by separating from them the circumstances of Time, and Place, and any other *Ideas*, that may determine them to this or that particular Existence.' Cf. II.xi.10-11 p.159-160; II.xii1-3 p.163-164. Cf. Berkeley, *Principles of Human Knowledge* ed. by Howard Robinson (Oxford 1996) Introduction §6-9 p.9-10.

[222] For a more extensive discussion of Locke's theory of abstraction and universals in relation to Hobbes and other early-modern philosophers, Cf. Ayers, *Locke I: Epistemology* 242-258.

[223] Locke is not entirely explicit as to whether universal and abstract ideas in general can exist without words or not. He suggests at *Essay* II.xi.10-11 p.159-160 that animals cannot construct these complex ideas because they lack the use of words. At the same time, he also maintains that humans without the capacity for language may nonetheless construct general ideas by the use of other kinds of general signs.

[224] Locke, *Essay* III.x.26-27 p. 505. Cf. IV.v.4 p.574-575; III.ii.6 p.407.

[225] Lenz, 'Locke as a Social Externalist', 60-62.

has to be amended. It turns out that Locke never explicitly discusses the function of words for complex ideas of substances, while he does provide a comprehensive account of the structural function and necessity of terms in the composition and retention of mixed modes. Locke thus states explicitly that without moral terms, the mind would not be able to retain the mixed modes used in moral reasoning:

> *The near relation* that there is *between Species, Essences, and* their *general Names*, at least in *mixed Modes*, will farther appear, when we consider, that it is the Name that seems to preserve those *Essences*, and give them their lasting duration. For the connexion between the loose parts of those complex *Ideas*, being made by the Mind, this union, which has no particular foundation in Nature, would cease again, were there not something that did, as it were, hold it together, and keep the parts from scattering. Though therefore it be the Mind that makes the Collection, 'tis the Name which is, as it were the Knot, that ties them fast together.[226]

Because the mixed modes of morality are not modelled on concrete objects in empirical reality, but entirely 'the Workmanship of the Mind,' they depend exclusively on words for stability.[227] Locke therefore suggests that in the case of mixed modes, the structural function of words is so important that without terminology, it is in fact impossible to retain moral concepts.[228] This conclusion is also supported by Locke's insistence that without words for numbers – which also classify as ideas of modes – humans are unable to count.[229] In addition, Locke maintains that the usual method of learning ideas of mixed modes is by the explication of the terms that stand for them.[230]

Locke thus clearly states that moral terms and mixed modes are mutually dependent. But does the fact that words have a structural role in the retention of ideas entail that the conventional signification of these words also influences or determines the composition of moral concepts? Considering this question in a recent paper, Benjamin Hill has maintained that there are indeed good reasons to ascribe this view to Locke. Thus some passages in the *Essay* suggest that Locke thinks that individuals model their moral concepts on the conventional signification of words:

---

[226] Locke, *Essay* III.v.10 p.435. Cf. III.v.11 p.435; III.v.15 p.437.

[227] Locke, *Essay* III.v.4 p.430.

[228] Locke, *Essay* III.v.11 p.435; II.xxii.4 p.289. Although Locke thinks that simple modes do not require words. Cf. Locke, *Essay* II.xviii.5-7 p.224-225.

[229] Locke, *Essay* II.xviii.5-7 p.224-225.

[230] Locke, *Essay* II.xxii.3 p.289.

> That the abstract *Ideas* of mixed Modes, being Men's voluntary Combinations of such a precise Collection of simple *Ideas*; and so the Essence of each Species, being made by Men alone, whereof we have no other sensible Standard, existing any where, but the Name it self, or the definition of that Name: We have nothing else to refer these our *Ideas* of mixed Modes to as a Standard, to which we would conform them, but the *Ideas* of those, who are thought to use those Names in their most proper Significations; and so as our *Ideas* conform, or differ from them, they pass for true or false.[231]

In this passage, Locke seems to claim that as mixed modes do not depend on any standard accessible to experience, their composition is derived from the conventional signification of the moral terms on which they depend. At the same time, Hill recognises that this single passage is not overwhelming textual evidence for this position, indicating either that Locke had not entirely made up his mind or found the issue not important enough to take a definite stance.[232]

Yet Hill nonetheless maintains that it makes sense to ascribe to Locke the view that the composition of mixed modes is determined or at least influenced by the conventional signification of moral terms. Hill argues that in order to function as moral concepts within Locke's epistemology, mixed modes would have to be composed in identical or at least similar ways by all members of a community. As collective norms of behaviour, morality only functions when everyone in society understands it in a similar way. And because the mixed modes of morality do not depend on any models in nature, it appears that the conventions of common use are the only standard by which individuals could effectively coordinate the composition of their moral concepts – and thereby their moral outlook. Upon this reading, the linguistic conventions specifying the signification of moral terms are thus indispensible to the actual function of the moral guidelines that enable sociability. Hill's interpretation thereby specifies what Locke means when he refers to language as 'the great Instrument and common Tye of Society.'[233]

Nonetheless, Hill's interpretation faces at least two objections. First of all, we have seen that Locke claims that the conventions of common use are neither

[231] Locke, *Essay* II.xxxii.12 p.388.

[232] Benjamin Hill, 'Locke on the Names of Modes', in: Margaret Cameron and Robert J. Stainton (eds.), *Linguistic Content: New Essays on the History of Philosophy of Language* (Oxford 2015) 176-200.

[233] Locke, *Essay* III.i.1 p.402

accurate nor precise. As mixed modes are not accessible to experience, this applies in particular to the conventional signification of moral terms.[234] Thus even if Locke thought that in practice the conventions of common use influence or determine the way we compose our moral concepts, he would have considered this an undesirable reality. And indeed, Locke urges his readers not to think merely with words, but rather to consider the ideas signified by them.[235] Secondly, in isolation, the view that the composition of mixed modes constitutive of moral reasoning depends on the conventional signification of terms would lead Locke towards a position of moral conventionalism. Unless the standard of common use exemplifies the universal morality of natural law the resulting conception of morality would be merely conventional. Although a singular interpretation has recently ascribed Locke a position of moral conventionalism, we will see below that this reading completely disregards Locke's belief in the existence of natural law – a universal morality decreed by God himself.[236]

Ultimately, Locke's remarks on the relationship between language and morality are not precise enough to draw any definitive conclusions. Thus Locke clearly suggests that moral concepts are regularly acquired through language. In addition, the conventional signification of moral concepts often functions as a standard for the composition of mixed modes to individuals. At the same time, Locke is not explicit on the extent of the influence of language on morality. Even though in practice we often acquire moral concepts through language, Locke suggests that we could – and, in light of the imprecision of linguistic convention, in fact we should – compose our mixed modes independent from the conventional signification of moral terms in order to arrive at a conception of natural law. In the following chapter, we will see that while Condillac borrowed much of his own philosophy of language from Locke, he took a more definitive

[234] Locke, *Essay* II.xxxii.12 p.388.

[235] Locke, *Essay* IV.v.4 p.575: 'Some confused or obscure Notions have served their turns; and many who talk very much of *Religion* and *Conscience*, of *church* and *Faith*, of *Power* and *Right*, of *Obstruction* and *Humours*, *Melancholy* and *Choler*, would, perhaps, have little left in their Thoughts and Meditations, if one should desire them to think only of the Things themselves, and lay by those Words, with which they so often confound others, and not seldom themselves.'

[236] Samuel Zinaich Jr., *John Locke's Moral Revolution: From Natural Law to Moral Relativism* (Lanham 2006) claims that when Locke declared that morality is capable of demonstration, he merely claimed that it would be demonstrable within a firmly established system of moral conventions. This interpretation of Locke as a moral relativist is shared by no other important commentator. For Locke's own view on moral convention, cf. *Essay* I.iii.21-22 p.81-82, where he describes the emergence of prejudice among moral conventions.

stance on this issue, concluding that the composition of our ideas, including those of morality, is inextricably tied to the conventions of language.

Despite some of the ambiguity in his position, Locke nonetheless makes an important contribution on the issue of the relation between language and morality. Ever since Aristotle, it had been commonplace for moral philosophers to associate man's propensity for sociability and morality with his capacity for language.[237] Locke's reflections on words explain even more thoroughly than Hobbes had done how this association between language and morality functions in practice. Thus Locke details how the use of words is necessary for the composition and transmission of moral concepts – and therefore for morality itself. In addition, Locke explains how the standard of common use requires individuals to compose their moral concepts in similar ways, thereby imparting a measure of agreement in their moral concepts. Finally, Locke's theory of signification also clarifies why despite the unifying role of common use, language necessarily remains an imperfect method of communication – especially in moral matters. Even though Locke not always reaches a definite conclusion on these issues, his discussion provided the impetus for many subsequent early-modern debates on language and its role in cognition and morality – of which a few will feature in subsequent chapters. Some scholars have therefore presented Locke's philosophy as the first linguistic turn in modern philosophy – and not entirely without reason.[238]

Mixed Modes and the Rule of Right

In the preceding, we have interpreted Locke as claiming that by reasoning with mixed modes, human beings may demonstrate morality, thereby ascertaining a conception of natural law. First we noted that by taking Locke's account at face value, these demonstrations turn out to be nothing more than the explication of moral concepts, yielding no more than is already contained in the definitions of moral terms. We therefore supposed that to understand how Locke thinks these demonstrations could yield veritable insight into morality – and in particular the morality of natural law – we would need to identify the method by which mixed modes are constructed. Subsequently, we saw that there are some grounds to ascribe Locke the view that the composition of mixed modes is determined or at

[237] Aristotle, *Politics* I.ii.1253a7.
[238] This is the argument of Losonsky, *Linguistic Turns in Modern Philosophy*.

least influenced by the conventional meaning of moral terms. Yet we concluded that while Locke's own statements on the issue are not conclusive, he suggests that even if words and their conventional meanings often influence our moral reasoning, this is ultimately not an ideal practice. In order to ascertain the universal morality of natural law, we should rather turn to a universal moral standard decreed by God. In the following, we will first inspect Locke's statements on the role of this moral standard in normative reasoning. Subsequently we will consider several interpretations on the question of how Locke thinks this standard may be ascertained.

In the passages on moral demonstration discussed above, Locke gives the impression that a conception of morality can be established by relating or comparing mixed modes to each other. Consequently, Locke seems to claim that the outcome of moral demonstration is dependent on the composition of these mixed modes themselves. Yet elsewhere in the *Essay*, Locke presents an alternative account of normative reasoning. In a chapter called *Moral Relations*, Locke seems to suggest that by themselves, mixed modes are morally neutral concepts that only acquire normative significance when compared with an already established rule of right:[239]

> But this is not all that concerns our Actions; it is not enough to have determined Ideas of them, and to know what Names belong to such and such Combinations of Ideas. We have a farther and greater Concernment, and that is, to know whether such Actions so made up, are morally good, or bad.[240]

Subsequently, Locke contends that the morality of these actions may be determined by establishing 'the Conformity, or Disagreement, Men's voluntary Actions have to a Rule.'[241] In view of these passages, we might conclude that after all, Locke did not think that mixed modes by themselves constitute morality.[242] Instead, mixed modes are nothing more than archetypes that should be

[239] Locke concedes in *Essay* II.xxviii.16 that in practice many moral concepts – such as *stealing* or *murder* – are not neutral terms, but already contain within themselves certain moral judgments. But Locke maintains that we should avoid this normativity in moral concepts themselves. Accurate moral judgment is then only possible if we establish the normativity of neutral and purely descriptive moral concepts with an external moral guideline.

[240] Locke, *Essay* II.xxviii.4 p.351.

[241] Locke, *Essay* II.xxviii.4.p.351.

[242] This is in fact the view of Forde, '"Mixed Modes" in John Locke's Moral and Political Philosophy', 600

compared to an external rule to determine their normativity. Moral demonstration would then turn out to be more akin to judicial procedure – the meticulous comparison of moral concepts to an already established law.

Locke subsequently discusses three types of moral rule by which the normativity of mixed modes may be established: the civil law, the rule of opinion, and the natural law decreed by God. Of these three moral guidelines, the first two are merely conventional and can only indicate conformity of actions to the laws or moral customs of a particular society. Locke maintains that 'The only true touchstone of *moral Rectitude*' is contained in the natural law decreed by our Creator.[243] This universal morality is a direct consequence of God's will, and indicates the intentions He has for mankind. Yet in the chapter *Moral Relations*, Locke does not explain how this natural law decreed by God may be ascertained. Of course God's intentions may be inferred through revelation. But Locke does not suppose that morality is solely discernable through the interpretation of scripture. As we have seen, he clearly thinks that natural law is also accessible to human reason. Above we have also seen that in other passages, Locke appears to claim that the absolute morality of natural law could be established through moral demonstrations with mixed modes. Yet in his discussion on moral relations, he rather suggests that reasoning with mixed modes only yields normative conclusions in reference to an already established rule. But if moral demonstrations with mixed modes by itself cannot produce knowledge of natural law, how is this universal moral guideline to be established?

This question has puzzled both modern scholars and Locke's early modern readers. Taking Locke's account of the foundations of morality at face value, David Hume accuses his predecessor of presenting a circular theory, as the normativity of mixed modes depends on a rule, while this rule itself is apparently constituted by reasoning with moral concepts.[244] Some modern scholars are inclined to agree with Hume that Locke's moral epistemology ultimately rests

[243] Locke, *Essay* II.xxviii.8 p. 352.

[244] David Hume, *An Enquiry Concerning the Principles of Morals* Appendix I §9 p.85: 'No, say you, the morality consists in the relation of actions to the rule of right; and they are denominated good or ill, according as they agree or disagree with it. What then is this rule of right? In what does it consist? How is it determined? By reason, you say, which examines the moral relations of actions. So that moral relations are determined by the comparison of actions to a rule. And that rule is determined by considering the moral relations of objects. Is not this fine reasoning?' Cf. Hume, *Treatise of Human Nature* III.i.1 p.293-302.

upon unstable or unclear foundations.[245] Others have attempted to clarify Locke's meaning and identify the basis for natural law by placing his account of moral demonstration in a broader context. In the scholarship of the past decades, two main lines of interpretation have emerged that attempt to salvage, explain and justify Locke's project of a demonstrative moral science by identifying an exterior standard that would lend normativity to moral concepts and demonstrations. The first line of interpretation claims that Locke considers human nature as the fundamental standard of morality, while the second presents Locke's moral philosophy as ultimately a form of consequentialism. In the following, we will investigate these lines of interpretation, starting with the view that Locke derives morality from a substantive conception of human nature, subsequently reviewing a consequentialist reading.

Human Nature as a Normative Standard

The first line of interpretation claims that in Locke's theory, the universal morality of natural law would have to be derived from certain substantial and general characteristics of human nature. This view is advanced by Hans Aarsleff, but argued most extensively by John Colman.[246] He suggests that we should turn to Locke's account of the state of nature in the *Two Treatises of Government,* which aim to provide concrete political solutions to the problems of the state of nature to understand how we might correctly compose moral concepts.[247] According to Colman, Locke's method for the definition of moral concepts would be a kind of conjectural history. This enquiry would establish man's character, needs and desires in the state of nature, and on the basis of these traits determine the proper moral concepts for natural man, stripped from all the particularities and conventions of society. These 'natural' moral concepts that concur with human nature could then be used in demonstrations to deduce natural morality.[248] Colman thereby ascribes to Locke's a method of moral epistemology somewhat similar to that of Hobbes.

---

[245] Mackie, *Problems from Locke* 208-209.

[246] Hans Aarsleff, 'The State of Nature and the Nature of Man in Locke', in: John W. Yolton, *John Locke: Problems and Perspectives* (Cambridge 1969) 99-136.

[247] Locke, *Two Treatises of Government* II.ii 103-106.

[248] John Colman, *John Locke's Moral Philosophy* (Edinburgh 1983) 180: 'As Locke's state of nature encapsulates his view of the fundamentals of human nature, it would seem that we should turn to his account of life in this state for an understanding of what human beings need to do in order to

Yet we may identify several objections to this interpretative approach. To begin with, it is doubtful whether Locke thinks that human nature actually has any universal character, desires, needs, or aims. Rather, Locke notes that individuals have widely diverging predilections, and consequently take pleasure in different kinds of experiences and activities. This disparity in aims and inclinations would preclude any attempt of defining a specific moral guideline that would lead individuals to a circumscribed notion of happiness. Accordingly, Locke agrees with his predecessor Hobbes that there is no such thing as a *summum bonum*, a universal notion of the Good, derived from a substantive conception of human nature, towards which all humans should strive to become virtuous and happy.[249] As Locke maintains that humans have widely diverging aims and desires, he thinks that they also disagree on what constitutes felicity. Locke therefore states that individual happiness is not attained by any specific aims or pursuits, but rather simply by the durable satisfaction of any desires we might have and the consequent achievement of pleasure.[250]

In response to this objection, Peter Schouls has maintained that despite Locke's rejection of the notion of a *summum bonum*, his account of human nature may nonetheless provide a solid foundation for natural law.[251] Schouls points out that Locke states that human beings are not only characterised by a desire for pleasure and aversion to pain, but also by free will and a capacity for reasoning.[252] According to Schouls, these fundamental properties of human nature would provide sufficient basis to construct a concrete theory of

---

achieve happiness. This in turn will show what properly belongs to morality, or what is the genuine content of our moral notions.'

[249] Locke, *Essay* II.xxi.55 p.269.

[250] Locke, *Essay* II.xxi.54 p.268-269.

[251] The main target of Schouls's argument is the interpretation of John Passmore, who claims that for Locke, human beings have no circumscribed nature at all, but are fully conditioned by their experiences. Passmore's interpretation is outlined in his 'The Malleability of Man in Eighteenth-Century Thought', in: Earl Wasserman (ed.), *Aspects of the Eighteenth Century* (Baltimore 1965) and idem, *The Perfectibility of Man* (London 1970).

[252] While Locke uses the term 'free will,' he does not deny that the will is causally determined. Not unlike Hobbes, Locke thereby takes a compatibilist position. Locke then often uses the term 'free will' to refer to man's ability to weigh competing desires, deliberating prudently about their foreseen consequences rather than merely acting on a whim. On Locke's compatibilism cf. Matthew Leisinger, 'Locke's arguments against freedom to will', *British Journal for the History of Philosophy* 25 (2017) 642-662. For Locke's views on liberty in practical deliberation cf. Tito Magri, 'Locke, Suspension of Desire and the Remote Good', *British Journal for the History of Philosophy* 8 (2000) 55-70.

morality.[253] Yet while Locke certainly considers these properties intrinsic to human nature, it is difficult to see how this narrow conception of human nature could in fact lead to any *concrete* normative conclusions. It may lead us to the conclusions that to act morally is to pursue our desires in a rational way. Yet beyond this conclusion, Locke's narrow conception of human nature tells us nothing about the question what it is to be rational – or by extension, what it is to be moral. Thus while Schouls provides a sensible interpretation of Locke's conception of human nature, it is puzzling why he assumes that this conception could provide the basis for any substantial normative theory.

Yet the most potent objection to the interpretations of Colman, Aarsleff, and Schouls originates in Locke's epistemology. Above we have seen that Locke claims that human understanding is unable to penetrate the real essences of empirical objects. By contrast, Locke envisions that in the case of ideas of mixed modes, we do have the ability to grasp both nominal and real essences. This is the reason why we are able to arrive at knowledge as incontestable as mathematics through moral demonstration. As any attempt to formulate a conception of human nature is based on empirical observation, the resulting conception will be a complex idea of substance. The inability to grasp the essences of empirical objects therefore also extends to any investigation of human nature. If we would nonetheless establish our moral reasoning on the basis of an empirically constituted conception of human nature, the desired certainty of our deductions will be compromised due to our inability to know the real essence of substances. Just as we cannot use empirical investigation to formulate exact solutions to mathematical problems also conducted by means of ideas of modes, reasoning with ideas of substances cannot provide any completely certain or accurate answers to ethical questions.

Indeed, Locke himself explicitly warns against using ideas of substances in moral reasoning. When defending his claim for the possibility of a moral science as certain as mathematics, Locke states:

> Nor let any one object, that the names of Substances are often to be made use of in Morality, as well as those of Modes, from which will arise Obscurity. For as to Substances, when concerned in moral Discourses, their divers Natures are not so

[253] Peter A. Schouls, *Reasoned Freedom, John Locke and Enlightenment* (Ithaca and London 1992) 63-72; Cf. Aarsleff, *The State of Nature and the Nature of Man* 116; John W. Yolton, *Locke and the Compass of Human Understanding* (Cambridge 1970) 170 for similar arguments.

> much enquir'd into, as supposed; *v.g.* when we say that *Man is subject to Law*: We mean nothing by *Man*, but a corporeal rational Creature: What the real Essence or other Qualities of that Creature are in this Case, is no way considered.[254]

In this passage, Locke explicitly states that the concept of 'man' to be used in moral demonstrations should not be a complex idea of substance that attempts – but will inevitably fail – to describe the essence of human nature. The idea of man to be used in moral deductions is a mixed mode – an archetype constructed by the mind independent from the facts of empirical reality.[255] For Locke, this archetype comprehends nothing more than the properties he attributes to human nature in the *Essay*: rationality, free will, and desire for pleasure and aversion to pain. And as we have seen above, these properties of human nature provide an insufficient basis for any substantial normative conclusions.

Ultimately, Locke's dichotomy between ideas of substances and ideas of modes is the most compelling reason why he cannot derive natural law from a conception of human nature. As we have seen, Locke's taxonomy of ideas classifies moral concepts together with mathematical notions as ideas of modes which are constructed by the mind at will. As mixed modes do not depend on any unknown real essences but are rather the mind's own constructions, Locke thinks that the mind has full insight into the composition and relations of these moral concepts it has itself created. This is then the main reason why Locke is convinced that like mathematics, reasoning with moral concepts is capable of achieving demonstrative certainty. Yet we have seen that by itself, Locke's account of mixed modes as ideas created by the mind would suggest that moral concepts are simply conventional. Furthermore, we have seen that Locke's own examples of moral demonstrations with mixed modes are ultimately nothing more than analytic explications of terms. To avoid the conclusion that mixed modes are either conventional or vacuous, Locke thus requires an external normative standard in which moral concepts could be modelled. Colman, Aarsleff, and Schouls, maintain that a conception of human nature would be able to function as this external normative standard. Yet we have seen that in the *Essay*, Locke himself prohibits this approach, as the introduction of a substantial

---

[254] Locke, *Essay* III.xi.16 p.516.

[255] John Simmons, *The Lockean Theory of Rights* (Princeton 1994) 24-25 agrees and recognises that the resulting moral concept would not be grounded in reality or indeed anywhere else – raising the question how such a conception of human nature could be justified.

conception of human nature, which is a complex idea of substance based on an unknown essence, into moral demonstrations would abrogate their demonstrative certainty. The interpretations of Colman, Aarsleff, and Schouls, therefore disregard Locke's distinction between ideas of substances and of modes.

Our combined investigation of Locke's epistemology and ethics has thereby also shown that the dichotomy between ideas of substances and of modes prohibits Locke from employing the approach, employed by Hobbes among others, of constructing a theory of natural law on the basis of a conception of human nature derived from empirical observation. As this dichotomy between ideas of substances and of modes precludes deriving normative conclusions from empirical facts, Locke thereby anticipates Hume's law that prohibits inferring values from facts.[256] Of course, unlike Hume, Locke does not develop this particular consequence of this dichotomy of ideas into an explicit formulation of the is/ought distinction. In fact, the interpretations by Colman, Aarsleff, and Schouls, illustrate that this dichotomy between facts and values is hardly a conspicuous element of Locke's epistemology. Nonetheless, our investigation of Locke's moral epistemology has shown that before Hume, Locke already arrived at the position that values cannot be derived from empirical facts. This is perhaps not surprising, as the is/ought dichotomy has been almost universally accepted among modern moral philosophers – at least among those working within the empiricist tradition and its successor of analytic philosophy.[257] Following Hume's law, these thinkers then attempt to identify basis of normativity that does not rely on any empirical facts. Yet in the remainder of this chapter, we will discover that Locke does not yet succeed in identifying a clear alternative source of normativity.

[256] Hume, *A Treatise of Human Nature* III.i.1 p. 293-302.

[257] Hilary Putnam, *Reason, Truth and History* 127-128 notes that the fact/value distinction has become such an institutionalised position in philosophy that he foresees that even his own critique of the dichotomy will likely fail to have considerable impact.

From Hedonism to Moral Consequentialism

Apart from the reading that Locke aimed to derive morality from a conception of human nature, a second major line of interpretation presents Locke's account of human motivation as the source of normativity in his moral epistemology. Like Hobbes before him, Locke maintains that all human actions and thoughts are ultimately motivated by the desire for pleasure and aversion of pain. A person without desires or aversions would lack any impetus for thought or action, thereby being reduced to a state of lethargy.[258] More importantly, Locke also follows Hobbes in his claim that on the level of the individual subject, the perception of good and evil is the consequence of our assessment of the propensity of things to cause pleasure or pain in us:

> Things then are Good or Evil, only in reference to Pleasure or Pain. That we call *Good*, which is *apt to cause or increase pleasure, or diminish Pain in us; or else to procure, or preserve us the possession of any other Good, or absence of any Evil*. And on the contrary we name that *Evil, which is apt to produce or increase any Pain, or diminish any Pleasure in us; or else to procure us any Evil, or deprive us of any Good*.[259]

Locke thereby clearly adopts a hedonistic position that, from the perspective of the individual at least, equates good to pleasurable and evil to painful experiences. Consequently, Locke maintains that '*Happiness* then in its full extent is the utmost Pleasure we are capable of, and *Misery* the utmost Pain.'[260]

Like Hobbes, Locke suggests that this happiness is may be pursued by carefully deliberating our actions.[261] According to Locke, we may evaluate our passions by trying to foresee the good or evil consequences of pursuing the objects of our desires. This ability to foresee the consequences of our actions in turn relies on our knowledge and capacity for reasoning.[262] Locke thereby

---

[258] Locke, *Essay* II.vii.3 p.129.

[259] Locke, *Essay* II.xxi.2 p.229.

[260] Locke, *Essay* II.xxi.42 p.258. Note the easily overlooked but significant difference with Hobbes's conception of happiness. Whereas Locke simply suggests that happiness is achieved by maximising pleasure, we have seen that Hobbes maintains that felicity consists of the durable satisfaction of successive and continually occurring desires over a period of time.

[261] Locke, *Essay* II.xxi.47-52 p.263-267.

[262] Locke appears to endorse an intellectualist position in *Essay* II.xxi.62 p.275: 'tis impossible any one should willingly put into his own draught any bitter ingredient, or leave out any thing in his power, that would tend to his satisfaction, and the compleating of his Happiness, but only by a *wrong Judgment*.' Although Locke adds that in practice, many people cannot be bothered to

appears to arrive at a theory of subjective normative judgment that resembles that the account of Hobbes we have encountered above in that it is both consequentialist and intellectualist. In view of Locke's hedonist conception of the Good and resulting consequentialism, some interpreters have therefore proposed that for Locke, the main purpose of morality is ultimately to facilitate individuals in their aim for happiness, conceived as the maximisation of pleasure. These scholars thereby interpret Locke as ultimately a consequentialist or proto-utilitarian thinker.[263] This interpretation is not only supported by Locke's hedonism and consequentialism, but has the added benefit that it does violate the dichotomy between facts and values that follows from Locke's epistemology.[264]

Nonetheless, this reading faces the difficulty that it is not as easily reconciled with Locke's adherence to the notion of morality as a natural law decreed by God. Locke not only believes that natural law has a Divine origin, but also that we are punished or rewarded for our obedience to God's law of nature in life after death.[265] For Locke, these Divine retributions and rewards are in line with his hedonistic presupposition – they are nothing more than the ultimate instances of pleasure and pain following our actions in life.[266] But Locke's belief

---

deliberate their actions attentively, as they are overwhelmed by a strong desire for transitory pleasures.

[263] Andrew Israelsen, 'God, Mixed Modes and Natural Law: An Intellectualist Interpretation of Locke's Moral Philosophy', *British Journal for the History of Philosophy* 21 (2013) 1111-1132; Leo Strauss, *Natural Right and History* (Chicago 1953) 248-251.

[264] Although, strictly speaking, the claim that all human thought and action is occasioned by the passions is of course already a statement about human nature, and therefore involves a complex idea of substance. Ultimately Locke's entire account of human understanding is in fact vulnerable to this objection. Locke's epistemology therefore appears to rely on the assumption, unarticulated by Locke himself but made explicit by his successors, that our perception of the operations of our own mind does not suffer from the epistemic limitations that plague regular empirical observation, and therefore provides us with certain rather than probable knowledge. For a critique of this assumption, see Rorty, *Philosophy and the Mirror of Nature* 155-163.

[265] Locke, *Essay* II.xxviii.8 p. 352. To be sure, Locke cannot definitively prove the reality of the afterlife and in this case we must be content with the assurance that its existence is highly probable. In light of Locke's claim that his moral science yields certain rather than merely probable conclusions, this would seem problematic. For how could we suppose that the conclusions of a science are certain if it is based on merely probable presuppositions? Steven Forde, 'Natural Law, Theology, and Morality in Locke', *Journal of Political Science* 45 (2001) 396-409, 400 notes that Locke takes a Pascalian position by assuming that the high probability of divine judgment is enough to compel us to follow God's law, simply because the small uncertainty over the reality of God's judgment is not compelling enough to incite us to disregard natural law; Cf. Locke, *Essay* II.xxi.44 p. 260-261.

[266] Locke, *Essay* II.xxviii.5 p. 351.

in Divine judgment also entails that the pleasures and pains of this life are only secondary to those experienced in the afterlife. Therefore the most important goal of morality is not merely to direct us to happiness in *this* life. While it does not necessarily exclude enjoyment of life on earth, natural law primarily serves to guide us towards pleasure in the eternal afterlife. Yet this raises questions about how the content of this natural law decreed by God may be ascertained. Is the pursuit of happiness in this life a reliable means of securing salvation in the afterlife? In other words, is the natural law decreed by God synonymous with a prudential rule that indicates the most durable and effective means of satisfying our desires during life on earth? Or does Locke think that salvation requires virtues other than the prudential pursuit of our desires?

Finding an answer to these questions is further complicated by the fact that Locke is generally interpreted as retaining a so-called *voluntarist* conception of natural law.[267] This voluntarism entails that for Locke natural law is a direct product of God's will, has an independent existence, and no necessary relation to the nature of things. Locke's position can be contrasted with that of *intellectualists* like Grotius or Leibniz, who held that natural law is a rational rule that can be derived from the nature of things or ascertained with the help of an innate and intuitive faculty or reason.[268] While they believed that God was ultimately responsible for this rational rule simply because He had created the universe, the intellectualists denied that natural law is a separate moral decree that God has imposed on an otherwise morally neutral universe.[269] On the basis of this position, Grotius famously stated that even if God would not exist, the precepts of natural law would still remain true because they are grounded in the nature of things rather than Divine will.[270] But for Locke, this position is unacceptable – and not only because it obviates God's role to natural law. Even if there would be a rational rule inherent in the nature of things, our epistemic

[267] Forde, *Locke, Science and Politics* 5-8.

[268] Perhaps rather confusingly, our current discussion employs the term 'intellectualism' in two different senses, denoting two distinct doctrines of moral philosophy both generally referred to by this term. Above our usage of the term 'intellectualism' referred to the view, most commonly associated with Socrates but also encountered in our chapter on Hobbes, that human beings do not act immoral willingly, but rather out of ignorance. Presently, we use the term 'intellectualism' to refer to a position on the status of natural law, which considers natural law as a rational rule that exists independently from Divine will.

[269] Cf. Schneewind, *Invention of Autonomy* 15-36 and 149-151 for a discussion of the distinction between voluntarism and intellectualism.

[270] Grotius, *On the Law of War and Peace* p.4.

limitations would make it inaccessible to us, as we are unable to know the real essences of substances. Consequently, our inability to grasp the true nature of things would make it impossible to use it as a basis for the definition an unquestionable moral guideline.[271] This reading of Locke's voluntarism therefore suggests that the natural law decreed by God need not be reflected in any properties of empirical reality. Like other voluntarist conceptions of natural law, Locke's theory therefore faces the objection that if morality is solely the consequence of God's will, it is not clear how human beings could ascertain the content of this law.[272] Accordingly, it becomes even more doubtful whether happiness in this life is a reliable indicator of the Divine decree that is enforced by the rewards and punishments of the afterlife.

Scholars in favour of a consequentialist reading of Locke's ethics have proposed several ways to neutralise this objection. Thus Leo Strauss has simply claimed that the inconsistencies that result from Locke's adoption of both hedonism and voluntarism 'are so obvious that they could not have escaped the notice of a man of his rank and sobriety.'[273] According to Strauss, Locke's claim that natural law is a Divine decree is insincere and merely intended to make his controversial hedonist doctrine palatable to his contemporaries by retaining the theological status of morality. A more promising line of interpretation that aims to resolve the tension between Locke's hedonism and his voluntarism is proposed by Israelsen and Tuckness among others, who maintain that while Locke indeed considered natural law a divine decree, his voluntarism does not necessarily entail that God could have devised natural law arbitrarily or independent from the rest of creation. To support this reading, these scholars claim that for Locke,

---

[271] Forde, '"Mixed Modes" in John Locke's Moral and Political Philosophy', 599 agrees with this assessment, while Israelsen, 'God, Mixed Modes and Natural Law: An Intellectualist Interpretation of Locke's Moral Philosophy' disputes it.

[272] Among others, this objection is made by Helvétius, *De l'homme* ed. by Jacques Moutaux (Paris 1989) IX.xv t.2 p.810: 'pour cacher l'absurdité de ce raisonnement, ils ajoutent que la justice d'en haut n'est pas celle de l'homme. Mais si la justice du Ciel est la vraie et que cette justice ne soit pas celle de la terre, l'homme vit donc dans l'ignorance de la justice. Il ne sait donc jamais si l'action qu'il croit équitable n'est point injuste, si le vol et l'assassinat ne sont point des vertus. Que deviennent alors les principes de la Loi naturelle et de la Morale ? Comment s'assurer de leur justesse et distinguer l'honnête homme du scélérat ?.'

[273] Strauss, *Natural Right and History* 220. For a more recent, though equally improbable argument that Locke's affirmations of belief in a Deity are insincere, and that he was in fact a closet atheist, see William T. Bluhm, 'Locke's Idea of God: Rational Truth of Political Myth?', *The Journal of Politics* 42 (1980) 414-438.

God is essential to natural law not because He is responsible for its promulgation, but rather because its enforcement requires His rewards and punishments in the afterlife.[274] Furthermore, Locke's belief in God's benevolence would preclude the possibility that God has prescribed humans a moral law that is not reflected in His own creation and would require humans to act in contradiction with their own nature.[275] This would suggest that despite his voluntarism, Locke thinks that certain salient properties of human nature reflect God's purpose for man's life on earth. Indeed, Locke sometimes tries to explain properties of human nature, such as the limitation of the mind to understanding the secondary qualities of empirical objects, as the consequence of God's conscious design.[276]

This particular reading of Locke's voluntarism would then allow for the reconciliation of Locke's hedonism with his conception of natural law as a divine decree. It suggests that Locke does not believe that God could have imposed an arbitrary natural law that would conflict with the natural abilities and inclinations of his creatures. Rather, His benevolence entails that there must be some level of conformity between human nature and God's moral law. Israelsen has therefore proposed that since God designed humans with the desire for pleasure and aversion of pain, these passions must be an indicator of His intentions for our life on earth. Thus while morality still has the metaphysical status of a natural law instituted by God, we need only inquire into the consequences of actions – the extent they cause pleasure or pain – to deduce their morality. Accordingly, Israelsen ascribes to Locke the view that natural law may be ascertained simply by maximising pleasure and avoiding pain.[277] Yet while Israelsen's consequentialist interpretation succeeds in identifying a possible source of normativity in Locke's theory, it is also contradicted by several of

[274] Israelsen, 'God, Mixed Modes and Natural Law: An Intellectualist Interpretation of Locke's Moral Philosophy', 1114; Alex Tuckness, 'The Coherence of a Mind: John Locke and the Law of Nature', *Journal of the History of Philosophy* 37 (1999) 73-90, 75-76 thus confirms that Locke is a 'ground voluntarist,' but denies that he is a 'content voluntarist.' This reading is supported by Francis Oakley, 'Locke, Natural Law, and God – Again', *History of Political Thought* 18 624-651, who shows that scholastic thinkers including Ockham retained a similar doctrine of voluntarism.

[275] Tuckness, 'The Coherence of a Mind: John Locke and the Law of Nature' 78.

[276] Martin Lenz, 'Locke's Lifeworld: The Teleological Role of Secondary Qualities', *History of Philosophy Quarterly* 36 (2019) 39-59 highlights that for Locke, God has consciously created humans with certain passions and epistemic limitations. According to Lenz, Locke sees these properties of human nature as reliable indicators of God's purpose for man during life on earth.

[277] Israelsen, 'God, Mixed Modes and Natural Law: An Intellectualist Interpretation of Locke's Moral Philosophy', 1128.

Locke's suggestions on how to achieve eternal bliss in the afterlife. Thus Locke maintains that the pleasures and pains of this life bear no proportion to those of the afterlife, and that we should aim to secure eternal bliss rather than transient pleasures.[278] Furthermore, Locke postulates that the practise of virtue does not always lead to happiness during life on earth, and is sometimes only truly rewarded in the afterlife.[279] Both these suggestions indicate that Locke in fact sees no necessary connection between man's worldly passions and Divine punishment and reward.

Locke thus appears to retain the view that natural law is a Divine decree that cannot simply be equated with a prudential guideline for the maximisation of pleasure during life on earth. Accordingly, it appears that Locke holds that not all desires and pursuits conform to God's intentions for humankind, and thereby to the law of nature He has decreed to guide us in our conduct. A recent interpretation by Elliot Rossiter in fact adopts this position as its premise in a renewed attempt to reconcile Locke's hedonist consequentialism with his notion of morality as a divine law. According to Rossiter, Locke indeed thinks that the pleasures and pains of this earth are designed by God to be indicative of those in life after death. Yet this does not necessarily entail that pleasures and pains are entirely equal. Thus Rossiter claims that Locke makes a distinction between transitory corporeal delights, and more lasting intellectual pleasures that accompany sociable conduct. According to Rossiter, this distinction between transitory delights and lasting pleasures functions as a signpost of divine intent. Locke would therefore hold that these sociable pleasures are a reliable indicator of God's intentions for us, and thereby of natural law.[280]

---

[278] Locke, *Essay* II.xxi.60 p.274: 'for since nothing of Pleasure and Pain in this Life, can bear any proportion to endless Happiness, or exquisite Misery of an immortal Soul hereafter, Actions in his Power will have their preference, not according to the transient Pleasure, or Pain that accompanies, or follows them here; but as they serve to secure that perfect durable Happiness hereafter.' Cf. II.xxi.44 p.261.

[279] Locke, *Essay* II.xxi.70 p.281: 'That a vertuous Life, with the certain expectation of everlasting Bliss, which may come, is to be preferred to a vicious one, with the fear of that dreadful state of Misery, which 'tis very possible may overtake the guilty; or at best the terrible uncertain hope of Annihilation. This is evidently so, though the vertuous Life here had nothing but Pain, and the vicious continual pleasure: which yet is for the most part quite otherwise, and wicked Men have not much the odds to brag of, even in their present possession.'

[280] Elliot Rossiter, 'Hedonism and Natural Law in Locke's Moral Philosophy', *Journal of the History of Philosophy* 54 (2016) 203-225.

While Rossiter's reading provides a novel perspective on the interpretation of Locke's moral philosophy, it also faces at least two objections. To begin with, Rossiter's reading is only marginally supported by textual evidence. Thus one of Rossiter's central claims, that Locke applies an essential distinction between transitory delights and intellectual or sociable pleasures, relies almost entirely on a brief remark in the *Ethica*, a fragmentary work that remained unpublished during Locke's lifetime.[281] The notion that these sociable pleasures are an indicator of natural law is then distilled from Locke's remark that God has ensured a connection between private virtues and public happiness.[282] Yet even in the *Ethica* itself, Locke concedes that depending on the circumstances, sociable and intellectual pursuits may not always lead to happiness, while doing evil may sometimes bring us pleasure. Locke claims that divine punishment and reward is necessary precisely because there appears to be no inherent connection between practising virtue and achieving happiness, conceived as 'the utmost Pleasure we are capable of.'[283] Furthermore, Rossiter's interpretation ascribes to Locke the view that not all desires are equally conducive to happiness. Yet above we have seen that Locke denies that there is any inherent distinction between various pursuits. Rather, Locke claims that due to different predilections, humans will have divergent desires and may attain happiness through diverse forms of pleasure. Accordingly, we have seen that

[281] Locke, *Ethica A* in: Mark Goldie (ed.), *Locke: Political Essays* (Cambridge 1997) 318: 'Pleasures are all of the mind, none of the body, but some consist in motions of the body, some in contemplations and satisfactions of the mind separate, abstract and independent from any motions or affections of the body. And these latter are both the greatest and more lasting. The former of these we will for shortness sake [call] pleasures of the senses, the other, pleasures of the soul, or rather, material and immaterial pleasures.'

[282] Locke, *Essay* I.iii.6 p.69: 'For God, having, by an inseparable connexion, joined *Virtue* and publick Happiness together; and made the Practice thereof, necessary to the preservation of Society, and visibly *beneficial* to all, with whom the Virtuous Man has to do; it is no wonder, that every one should, not only allow, but recommend, and magnifie those Rules to others, from whose observance of them, he is sure to reap Advantage to himself.' Yet by itself, this position does not entail or even imply that the content God's moral law can be equated to the maximisation of pleasure on earth for the *individual* agent.

[283] Locke, *Ethica A* 319: 'Happiness therefore is annexed to our loving others and to our doing our duty, to acts of love and charity, or he that will deny it be so here because everyone observes not this rule of universal love and charity, he brings in a necessity of another life (wherein God may put a distinction between those that did good and suffered and those who did evil and enjoyed by their different treatment there) and so enforces morality the stronger, laying a necessity on God's justice by his rewards and punishments, to make the good the gainers, the wicked losers.' Schneewind, *The Invention of Autonomy* 158 agrees with this assessment.

Locke rejects the notion of a *summum bonum*, or a universal good to which all humans should aim to attain happiness. By contrast, Rossiter claims that Locke does make an inherent distinction between desires in view of their propensity to lead us to happiness. His interpretation would therefore also ascribe to Locke a notion of a universal good that is clearly rejected in the *Essay*.

In the end, neither Rossiter nor Israelsen thus entirely succeeds in their attempt to reconcile Locke's hedonism with his conception of morality as a divine decree. On the one hand, Locke claims that as all our thoughts and actions are occasioned by the desire for pleasure and aversion of pain, happiness is nothing but the utmost pleasure we may achieve. Yet on the other hand, we have seen that Locke clearly does not believe that there is a necessary connection between the pleasures of this life and the moral rule decreed by God and enforced in the afterlife.[284] Despite his hedonism, Locke therefore continues to hold that true virtue is something other than mere prudence, and that the most important consequences of following or trespassing natural law are only experienced in the afterlife. Yet while there are legitimate reasons to dispute the interpretations of both Rossiter and Israelsen, their contributions are nonetheless valuable because they highlight the proximity of Locke's moral philosophy to a more consistently consequentialist and utilitarian position. When stripped of its eschatological element, Locke's moral philosophy clearly classifies as a consequentialist theory. For this reason, Locke is often identified by scholars as one of the major forerunners of the utilitarian moral theory developed during the latter half of the eighteenth century by Bentham and Helvétius among others.[285]

Conclusion

In the end, we must conclude that it is not feasible to present an entirely coherent and convincing interpretation of Locke's moral epistemology if we intend to do justice to all of his philosophical commitments and presuppositions. It would therefore appear that ultimately, Locke does not entirely succeed in demonstrating his claim that '*Morality is capable of Demonstration*, as well as Mathematicks.' Indeed, it appears that later in life Locke himself became less optimistic about the possibilities of establishing morality through reasoning alone. Thus in the *Reasonableness of Christianity*, published some five years after

[284] Ayers, *Locke II: Ontology* 198-199 agrees with this assessment.
[285] Heydt, 'Utilitarianism before Bentham.'

the *Essay*, Locke wrote that it may be 'too hard a task for unassisted reason to establish morality, in all its parts, upon its true foundations, with a clear and convincing light.'[286] Yet Locke's self-confessed failure to provide a coherent moral philosophy should not deter us from studying it. In fact, the problems in Locke's moral epistemology highlighted in our investigation reveal some of the most important questions and challenges faced by all early-modern moral philosophers who rejected both scholastic essentialism and the Cartesian notion of innate ideas, and came to embrace an empiricist account of human understanding.

Throughout our discussion, we have seen that many of the perplexities of Locke's moral philosophy are then ultimately a reverberation of his own theory of knowledge. Previously, natural law theorists had either supported their conception of morality with a version of Scholastic essentialism, or they had argued that human beings simply have an innate faculty of normative judgment. Yet we have seen that Locke dismisses both approaches to grounding morality, as his empiricism rejects both essentialism and any form of innatism. Furthermore, Locke rejects the approach, applied by Hobbes among others, of deriving morality from a substantive conception of human nature. We have seen that for Locke, this approach is unsound because it violates the dichotomy between ideas of substances and moral concepts that is in integral part of his epistemology. Accordingly, we have established that Locke already proposes a distinction between facts and values that most commonly associated with David Hume.

As Locke's epistemology prohibits many of the strategies of grounding morality employed by his predecessors, his philosophy is left with only three possible sources of normativity. Firstly, Locke recognises that morality is a social construct transmitted by and contained within conventional language. Secondly, true to his empiricist principles, Locke binds morality to a normative component of experience – the feeling of pleasure and pain. Thirdly, Locke at the same time remains attached to the notion that natural law is not mere prudence, but rather a divine decree enforced by the punishments and rewards of the afterlife. As Locke is convinced that mere convention is inadequate as a source for morality,

[286] Locke, *The Reasonableness of Christianity* in: *Works* VII (London 1823) p.139. Cf. Andrea Sangiacomo, 'Locke and Spinoza on the epistemic and motivational weaknesses of reason: the Reasonableness of Christianity and the Theological-Political Treatise', *Intellectual History Review* 26 (2016) 477-495. Locke voiced a similar sentiment in a letter his friend Molyneux. Cf. Colman, *John Locke's Moral Philosophy* 138; Wolterstorff, *John Locke* 141-142.

he is left with the latter two sources of normativity. Yet we have seen that Locke glosses over the question how these two disparate foundations of morality relate to each other in his moral philosophy. Throughout our discussion, we have encountered many scholarly efforts to reconcile Locke's hedonism with his eschatology. Yet rather than aiming to provide a fully coherent interpretation, our own reading has sought to highlight the fissures and tensions on Locke's moral philosophy. Accordingly, we may conclude that Locke's consequentialism and hedonistic account of human motivation clearly prefigures the utilitarian conception of morality that would be developed during the eighteenth century by thinkers like Helvétius. At the same time, we must recognise that Locke holds fast to an older notion of virtue, according to which morality is not mere prudence, but rather a moral law ultimately derived from God. By diagnosing rather than resolving this fundamental tension in Locke's thought, we may recognise that his theory as representative of a crucial juncture in the history of moral philosophy, as it exemplifies the transformation of a teleological conception of virtue towards a purely prudential notion of morality.

3. Language and Experience: Condillac's Conception of Human Nature and Morality

Introduction

John Locke's *Essay concerning Human Understanding* was extremely influential throughout the eighteenth century – especially so among the French *philosophes*.[287] Among these French thinkers, the Abbé Étienne Bonnot de Condillac (1714-1780) was clearly Locke's most dedicated follower. Yet while his thought has sometimes been considered little more than 'a supplement to Mr. Locke's *Essay*,' Condillac in fact devised his own philosophy, offering new perspectives while accommodating it to contemporary debates.[288] To begin with, Condillac devised an original theory of human nature, which is particular notable for its conception of the mind as shaped by habit and structured through language. Condillac thereby expels any remaining notion of innate properties from his theory of human understanding. Instead, he claims that man's cognitive faculties are nothing but habits acquired by means of experience. Condillac thereby conceives human nature as even more flexible and dependent on external factors than his predecessors had done.

In addition, Condillac claims that the more advanced abilities of human understanding, including scientific and normative reasoning, can only be developed through the use of language. To be sure, we have seen that Hobbes had made a similar point. In addition, Locke's philosophy of language equally suggests a close connection between language and thought – especially in moral matters. Yet Condillac expands upon their position by concluding explicitly that due to the inextricable connection between language and thought, human reasoning is determined by the particular composition of language. With respect to morality, this would entail that human beings compose moral concepts on the basis of the conventional signification of moral terms. Consequently, the normative reasoning of individuals would reflect the conventions of the language

[287] On Locke's influence in France see Jorn Schosler, *John Locke et les Philosophes francais. La critique des idées innées en France au dix-huitième siècle* (Oxford 1997).

[288] The English translation of Condillac's *Essay on the Origin of Human Knowledge* was first published in 1756 with this subtitle. Cf. Hans Aarsleff, 'introduction', in: Condillac, *Essay on the Origin of Human Knowledge* ed. and trans. by Aarsleff (Cambridge 2001) xviii.

prevalent in their society. Condillac therefore concludes that the progress of knowledge – including knowledge of morality – requires the wholesale reform of language itself.

In subsequent chapters we will see that this conception of human nature as flexible and shaped through language also became influential among other French *philosophes* more exclusively focused on ethical questions.[289] Thus Helvétius's theory, discussed in our next chapter, that human nature is entirely determined by experience was clearly inspired by Condillac's philosophy. Furthermore, reflecting Condillac's proposal for the perfection of language, Helvétius also maintained that moral reform could not succeed without revising and consolidating the conventional signification of moral terms. In a different manner, Condillac's views on language and its role in the dissemination of ideas – and in particular ideas of morality – clearly inspired Rousseau's reflections on language and the transformation of natural man into a civilised being. In view of this influence on contemporary *philosophes*, a major purpose of our present discussion of Condillac will be to provide a background to our subsequent exploration of Helvétius's and Rousseau's philosophies. As they were primarily interested in morality rather than epistemology, Helvétius and Rousseau borrowed many elements from their account of human understanding from Condillac. Especially our subsequent investigation of Rousseau's conception of human nature will therefore benefit from an understanding of Condillac's theory of the mind as shaped by experience and language.

Yet apart from providing a context to our subsequent discussions, the present chapter will also aim to present an original interpretation of Condillac's philosophy. We will thereby apply the same approach as in previous chapters and inquire how Condillac thinks human beings may understand morality by means of reasoning. Condillac maintains that the question how to acquire true and accurate ideas can only be answered by understanding how we acquire ideas in

[289] Condillac's fellow *philosophes* acknowledged the importance of his philosophy to the development of contemporary thought. Thus Rousseau ranked Condillac among 'les meilleurs raisonneurs et les plus profonds métaphysiciens.' Cf. Ellen Mcnive Hine, *A Critical Study of Condillac's* Traité des systèmes (The Hague 1979) 1. This high estimation of Condillac was shared by Voltaire among others. See: Voltaire, *Correspondance* ed. T. Besterman (Geneva 1962) 14319, October 1768. Apart from his influence on the *philosophes* of the eighteenth century, Condillac's philosophy of language also inspired the *ideologues* of the early nineteenth century. Cf. Cheryl Welch, *Liberty and Utility: the French Ideologues and the Transformation of Liberalism* (New York 1984) 51-54.

the first place – in other words, by studying the mind and its operations. As Locke before him, Condillac thereby derives normative conclusions about how to reason from a descriptive account of the operations of the mind.[290] Furthermore, Condillac does not consider reason as some innate faculty, but rather as 'nothing other than our knowledge of the manner in which we must govern the operations of the soul.'[291] In order to discover how Condillac believes we may ascertain morality, we will thus have to ascertain how he thinks human beings should conduct their mental operations. We will then first reconstruct Condillac's theory of human understanding as developed by means of experience and language. Following an exploration of Condillac's conception of human nature, we will then be in the position to reconstruct his account of moral judgment. We will find that while Condillac provides an account of moral judgment that on some points diverges from that of Hobbes and Locke, his moral philosophy as a whole declines to confront several ethical questions that had seemed crucial to his predecessors. Finally, during our investigation of Condillac's account of human understanding, we will find that he thinks that human knowledge has been disfigured by prejudices and unsound cognitive habits disseminated in society by means of language. We will then discover why Condillac maintains that the advance of judgment in both empirical science and morality requires the wholesale reform of both education and the conventions of language.

Expanding and Revising Locke's Theory of Human Understanding

Condillac himself acknowledges that John Locke's *Essay* is the primary source of inspiration for his theory of human understanding.[292] Even if many of Condillac's conclusions ultimately diverge from those of his illustrious predecessor, he acquires his fundamental presuppositions from Locke's philosophy. Thus Condillac firmly rejects the concept of innate ideas, noting that Locke may in fact have spent too much time and effort on refuting the doctrine.[293] Condillac also agrees with Locke that all our ideas are therefore derived or composed from

[290] Condillac, *Essay on the Origin of Human Knowledge* ed. and trans. by Hans Aarsleff (Cambridge 2001) Introduction p.5: 'We must never forget that our first aim is the study of the human mind, not to discover its nature, but to know its operations, to observe how artfully they interact, and how we ought to conduct them in order to acquire all the knowledge of which we are capable.'

[291] Condillac, *Essay* I.ii.92 p.63.

[292] Condillac, *Essay* Introduction p.4.

[293] Condillac, *Essay* Introduction p.7.

simple ideas acquired through either sensation or reflection.[294] In addition, Condillac adheres to Locke's essence agnosticism – the claim that the real essences of substances are inaccessible to human perception.[295] Although Condillac does not stress this doctrine as much as Locke had done – possibly because by the middle of the eighteenth century its acceptance no longer required elaborate argumentation. Finally, Condillac largely adopts Locke's classification of ideas, subdividing them into complex ideas of substances and of modes, both composed of simple ideas acquired through experience. Yet Condillac appears to discard Locke's category of ideas of relation, as he declines to mention it at all.

Yet while Condillac follows Locke in his most fundamental presuppositions, he does reject an important assumption of his English predecessor. In our previous chapter, we have seen that Locke tacitly assumes that the mind naturally has the ability to compose, recall and relate the ideas it has acquired.[296] According to Condillac, Locke thereby relies on the unsubstantiated position that man's cognitive abilities are somehow innate to the mind. Alternatively, Condillac maintains that these cognitive abilities are nothing but habits acquired through experience:

> Thus this philosopher [Locke] is content to recognize that the mind perceives, thinks, doubts, believes, reasons, knows, wills, reflects; that we are convinced of the existence of these operations because we find them in ourselves, and that they contribute to the progress of our knowledge; but he did not suspect that they could be only acquired habits; he seems to have regarded them as something innate, and he says only that they are perfected through use.[297]

---

[294] Condillac affirms in the *Essay* I.i.4-5 p. 11-12 that the mind's perception of its own operations constitutes a second source of ideas besides sensation. Yet rather confusingly, Condillac does not use the term 'reflection' to refer to this form of perception as Locke had done. Instead, Condillac adopts the commonplace conception of reflection as a form of introspective thought (Cf. *Essay* I.ii.5 §48 p.41). However, later in his career, such as in the 'Extrait raisonnée', in: idem, *Traité des sensations augmenté de l'extrait raisonnée* (Paris 1984) 240, Condillac decided to dispense with Locke's notion of reflection, leaving sensation as the only remaining source of ideas. Cf. Cassirer, *The Philosophy of the Enlightenment* 199-200.

[295] Condillac, *Treatise on Sensations* in: *Philosophical Writings of Etienne Bonnot, Abbé de Condillac* trans. by Franklin Philip and Harlan Lane (Hillsdale 1982) Introduction p.168; *Logic* in: *Philosophical Writings* trans. by Philip and Lane I.vi p.361; II.i p.384.

[296] Locke, *Essay* II.x-xi p.149-163.

[297] Condillac, *Treatise on Sensations* Introduction p.159; Cf. *Essay* Introduction p.7; II.ii.3 §39 p.214; I.ii.8 §74 p.52.

Our cognitive faculties are thus not properties natural to the understanding that merely need to be perfected through exercise. Instead, Condillac thinks that the abilities to compose, decompose, and relate ideas are in fact nothing more than habits, acquired gradually by experience. This position leads Condillac to the conclusion that: 'Judgment, reflection, desires, passions, and so forth are only sensation itself differently transformed.'[298] Condillac thereby adheres uncompromisingly to the metaphor of the mind as a *tabula rasa*. When Locke described the mind as a blank slate at birth, he merely intended to portray it as devoid of ideas. Condillac takes the metaphor more literally and claims the mind naturally contains neither ideas nor the cognitive faculties to process ideas received through sensations.

In order to substantiate this claim, Condillac describes at length how an individual may acquire the cognitive faculties peculiar to human beings. As we will see, this account shows some surprising similarities to Hobbes's theory of human cognition as outlined in the first chapter – an affinity that is acknowledged neither by Condillac himself nor by any modern interpreters.[299] Like Hobbes, Condillac thinks that besides the capacity for sense-perception, the mind is naturally endowed with three fundamental properties. The first is reminiscence – the ability to store sensations and ideas for later recollection. Yet while it enables the retention of ideas, by itself reminiscence does not allow the mind to recall stored sensations at will and may only recall previous sensations when present experience provides occasion for recollection.[300] Secondly, Condillac claims the mind naturally has the capacity to discern the difference and similarity of ideas, giving rise to the ability called judgment.[301] Finally, like Hobbes and Locke before him, Condillac thinks sensations are always intermixed

[298] Condillac, *Treatise on Sensations* Introduction p.171.; Cf. VII p.202.

[299] For a general overview of the influence of Hobbes's philosophy in France, see: Robin Douglass, *Hobbes and Rousseau: Nature, Free Will, and the Passions* (Oxford 2015) 21-61, which however declines to mention either Condillac or Helvétius as thinkers who incorporated elements from Hobbes's thought. Israel, *Enlightenment Contested* 228-230 in fact argues that no major thinker except Bayle and Rousseau engaged meaningfully with Hobbes's philosophy.

[300] Condillac, *Essay* I.ii.2 §15 p.25; I.ii.4 §37 p.36.

[301] Condillac, *Essay* I.ii.8 §69 p.51.

with feelings of pleasure and pain. According to Condillac, the desire for pleasure and aversion to pain are then the impetus to all thought and action.[302]

Like Hobbes before him, Condillac maintains that these three elemental abilities of the mind – reminiscence, judgment and the feeling of pleasure and pain – allow for the development of basic cognitive functions. In his *Treatise on Sensations*, Condillac further demonstrates his theory by conducting a thought-experiment, imagining the creation of an animated statue that is somehow endowed with these same elemental properties of the human mind. Condillac then conjectures what would happen if the senses of the statue would be gradually enabled. Thus he envisions that the statue is roused into action as soon as its sense of smell is activated. Immediately the statue will start to seek out pleasant smells, while avoiding disagreeable odours. As reminiscence will enable it to recall previous sensations, the statue will gradually learn how and where to satisfy its desire for pleasurable scents. Subsequently, Condillac describes how the gradual activation of the other senses incites the statue to search for other types of pleasurable sensations. Finally, it is the sense of touch that makes the statue realise that its sensations are actually occasioned by objects external to itself.[303]

Following the cognitive development as outlined the *Treatise on Sensations*, the statue will have desires, a will, some form of judgment and it will probably able to take care of its own preservation. But the thought experiment of the animated statue clearly does not describe the development of the more advanced cognitive abilities of human beings. Accordingly, Condillac has sometimes been accused of providing an incomplete account of human understanding. As Condillac claims that the mind is entirely devoid of innate properties and solely conditioned by experience, his account in the *Treatise on Sensations* would present the mind as merely a passive entity and would

[302] Condillac, *Treatise on Sensations* Introduction p.171.; Cf. VII p.202; *Essay* I.ii.11 §106 p.69: 'the influence of the passions is so great that without it the understanding is virtually at a standstill, so much so that for lack of passions there is barely any intellect left.'

[303] Cassirer, *The Philosophy of the Enlightenment* 118 explains that the *Treatise on Sensations* was probably intended as a reply to Diderot's *Letter on the Blind*, which suggested that Condillac's philosophy as outlined in the *Essay on the Origin of Human Knowledge* could be interpreted as a form of Berkeleyan idealism. Condillac's response to Diderot's interpretation was that through the sense of touch we are able to assure ourselves of the fact that external objects do exist. Cf. Dale Jacquette, 'Condillac's Analytic Dilemma', *History of Philosophy Quarterly* 32 (2015) 141-159 for a modern interpretation that portrays Condillac as an unintentional idealist.

therefore be insufficient to account for the complexity and variety of human thought and experience.[304] Yet this critique fails to appreciate that Condillac did not intend his thought experiment of the animated statue as an account of the full development of human cognition. Condillac explicitly states that the statue is limited to what he calls 'practical knowledge,' and lacks the means for theoretical investigation.[305] The statue therefore possesses what Condillac refers to as 'instinct' and is therefore more comparable to an animal than to a human being.[306] The statue is not yet capable of abstract or universal reasoning, nor is it a social being that can live together with others.[307] Like Thomas Hobbes before him, Condillac thinks that the development of both abilities depend entirely on the use of language.

Words and Ideas

Condillac describes three ways by which more advanced forms of thought are made possible by connecting ideas to linguistic signs. First of all, Condillac reiterates Locke's suggestion that words have an important structural function in the composition of ideas. Especially in the case of more complex aggregates of ideas, such as universals, abstractions and moral concepts, linguistic signs are required to make these mental constructions stable and accessible to the mind:

> The mind is so limited that it cannot recall a large quantity of ideas so as to make them the object of reflection all at the same time. Nevertheless the mind must often consider several of them together. It does that with the help of signs which, by uniting them, makes it possible for it to regard them as if they were a single idea.[308]

[304] Gaukroger, *The Collapse of Mechanism and the Rise of Sensibility* 414.

[305] Condillac, *Treatise on Sensations* IV.i p.307.

[306] Condillac, *Essay* I.ii.4 §43 p.39 defines instinct as: 'An imagination which in the presence of an object revives the perceptions that are immediately connected with it and which by that means guides all kinds of animals without the assistance of reflection.' Cf. Condillac, *Traité des animaux* ed. by M. Malherbe (Paris 2004) II.i-ii p.149-154.

[307] Hans Aarsleff, 'The Tradition of Condillac: The Problem of the Origin of Language in the Eighteenth Century and the Debate in the Berlin Academy before Herder', in: idem, *From Locke to Saussure, Essays on the Study of Language and Intellectual History* (London 1982) 146-209, 154; Aarsleff, 'Condillac's Speechless Statue', in: idem, *From Locke to Saussure* 210-224, 214. Wanda Wojciechowska, 'Le sensualisme de Condillac', *Revue Philosophique de la France et de l'Étranger* 158 (1968) 297-320 also challenges the interpretation of Condillac as a simplistic sensationist, although her interpretation attributes a greater attachment to innatism to Condillac than warranted by textual evidence.

[308] Condillac, *Essay* I.iv.1 §6 p.80. Cf. I.iv.2 §27 p.90-91.

Linguistic signs thus provide unity to complex notions constructed from otherwise disparate simple ideas. As we have seen in our previous chapter, Locke had maintained that this structural function of words was especially crucial in the case of mixed modes. Yet while Condillac agrees that moral ideas cannot be conceived without signs, he claims that the same applies to complex ideas of substances. Criticising Locke for failing to appreciate this last point, Condillac states that all complex ideas – both purely intellectual concepts as well as ideas derived from experience – depend on signs for stability and accessibility to the understanding.[309] Condillac thereby takes a position very similar in practice – although not identical in theory – to Hobbes's nominalism regarding universals and abstractions.[310]

Secondly, Condillac claims that by connecting ideas to conventional signs, the mind acquires control over its own operations. Above we have seen that without the use of language, the animated statue cannot recall and inspect ideas at will, but only recollect ideas and formulate judgments when experience provides occasion. In this sense, the mind of the statue as described in the *Treatise on Sensations* appears to be a passive entity. While it certainly has some ideas, and it is motivated by a disposition to attain pleasure, the statue does not yet have the ability to recollect, scrutinise and connect ideas at will – at least not without any concrete external stimulus. Condillac maintains that to attain this ability, the statue will have to acquire the use of linguistic signs. Condillac thus concludes that it is only through signs that a person becomes able to direct its own attention and to consider its ideas as it pleases.[311] Condillac's conclusion is very similar to Hobbes's view that it is only by means of signs that the mind may

[309] Condillac, *Essay* I.iv.2 §27 p.90-91; I.iv.1 §7-8 p.81.

[310] We have seen that Hobbes's nominalism follows from his view that abstract or universal *ideas* are impossible to conceive. Consequently, abstractions and universals exist in the form of linguistic terms with a definition specifying their reference. By contrast, Condillac does not deny that universal and abstract ideas can exist. He merely says that they cannot be composed or retained without being attached to a linguistic sign. Nonetheless, we will see below that for Condillac there is generally such a strong connection between sign and idea that in practice the two can hardly be separated.

[311] Condillac, *Essay* I.ii.5 §47 p.41: 'As soon as memory is formed and the exercise of the imagination is within our power, the mind begins to be set free from its former dependence on the objects that acted on it, owing to the signs that memory recalls and the ideas that imagination revives. With the full ability to recall things it has seen, it can direct its attention to them, away from those it has before the eyes at the moment.' Cf. I.ii.4 §37-39 p.36-37.

direct its attention beyond present sensations, thereby acquiring the ability for independent thought.

Finally, Condillac claims that it is only through signs that the mind may establish relations between ideas. These relations enable the mind to associate and compare ideas at will – thereby facilitating reasoning. In the previous chapter, we have seen that Locke maintained that these relations between ideas are themselves registered by a distinct type of ideas – so-called ideas of relation. Yet Condillac discards this category of ideas and assumes that by themselves, ideas are nothing but isolated collections of memorised sense impressions. Instead, Condillac maintains that 'Ideas connect with signs, and it is, as I will show, only by this means that they connect among themselves.'[312] Condillac thus maintains that it is only by learning a language – or more accurately, by acquiring the use of signs – that the mind gains the ability to establish and perceive relations between ideas. Again, we have seen in our first chapter that Hobbes had also maintained that the ability of the mind to establish relations between ideas depends on the use of language.[313]

Following his revision of Locke's philosophy, Condillac thus accords language a much more prominent role in his theory of the understanding. In fact, language has such a crucial function in Condillac's theory of epistemology that it becomes largely responsible for the organisation of the mind. Above we have seen that Condillac denies that the mind has any innate structure and claims that its cognitive faculties are nothing but habits acquired through experience. Yet it turns out that especially in the development of the higher cognitive abilities,

[312] Condillac, *Essay* Introduction p.5. Cf. I.ii.107 p.69. Cf. François Duchesneau, 'Condillac et le principe de liaison des idées', *Revue de Métaphysique et de Morale* 1 (1999) 53-79.

[313] In fact, Condillac's position omits an important element of Hobbes's account. In our first chapter we have seen that Hobbes maintains that the use of signs allows the mind to establish relations between ideas *at will*. Yet according to Hobbes, humans and animals alike may already establish relations of causality between successive objects of experience even before the acquisition of language. This ability, later described by David Hume as an instinctive habit, then allows for a prudential understanding of causal relations, enabling both animals and humans without the use of words to ensure their own preservation. It appears that Condillac's thought experiment in the *Treatise on Sensations* ascribes a similar ability to the sensitive statue. Nonetheless, Condillac states unequivocally that the mind may only establish relations between ideas by means of signs. Possibly, Condillac thereby merely means, with Hobbes, that the mind may only establish relations between ideas *at will*. Yet it is more likely that the discrepancy follows from the limitations of the Lockean theory of causality adopted by Condillac and later dissected by Hume in his *Enquiry concerning Human Understanding* IV p.25-39 which does not provide a clear explanation of the method by which the mind establishes causal relations between objects of experience.

these habits rely on the cognitive functions of language. Thus language provides the mind with a network of signs that allow it to structure, memorise and connect its ideas, thereby enabling higher cognitive faculties including reason. Rejecting the assumption that the mind has an innate structure or natural faculties, Condillac presents language as an alternative mechanism organising the mind. Thus while Condillac claims the faculties of the mind are accumulated gradually through experience, they are facilitated and expressed only with the help of the linguistic signs that comprise a language.

Linguistic Relativism?

Evoking Hobbes's views on the cognitive use of language, Condillac thus conceives an inextricable connection between words and ideas, stating that 'the art of reasoning reduces to the art of speaking well. To speak, to reason, to formulate general or abstract ideas, are thus fundamentally the same thing.'[314] Condillac even maintains that language not only determines our reasoning, it even regulates the imagination – and as such influences our creative and artistic expression.[315] Yet even if Condillac follows Hobbes in his theory that the use of signs is crucial to the development of human understanding, he draws different conclusions from this position. In our first chapter we have seen that Hobbes showed only limited interest in the question whether, in view of the cognitive function of signs, language may not only enable but also influence or even determine our thoughts. To be sure, Hobbes was concerned that persons are often swayed by rhetorical or insignificant speech. Yet we have established above that Hobbes did not present any substantial account of how the signification of terms is regulated by a standard of common use. Accordingly, we concluded that Hobbes did not suppose that the ideas in the minds of individuals may be determined by the conventions of language.

Condillac, on the other hand, does maintain that language may influence or determine the composition of our ideas. He arrives at this conclusion by combining an approximation of Hobbes's position on the relation between words and ideas with Locke's reflections on the role of common use in regulating

[314] Condillac, *Logic* II.v p.398. Cf. Condillac, *Logic* II.iv p.395: 'We think with them. Rulers of our judgment, they determine our knowledge, opinions, and prejudices. In short, they do in this domain everything good or bad. Such is their influence, and it could not have happened differently;' Aarsleff, 'The Tradition of Condillac', 165.

[315] Condillac, *Essay* II.i.15 §147-156 p.187-192

signification. Following Locke, Condillac presents language as a set of socially constructed conventions upheld by tacit consent. As Condillac considers language an inherently communal institution, it is non-existent in completely solitary individuals, who Condillac considers comparable both socially and intellectually to the statue-man described in the *Treatise on Sensations*.[316] According to Condillac, language is thus not only the occasion for, but also the product of social intercourse.[317] For Condillac, as for Locke, each language is thus a particular socially constructed artefact. In our previous chapter we have seen that while Locke was sensitive to the conventional nature of language, he avoids explicit linguistic relativism by supposing a more diffuse connection between words and ideas. Thus Locke seems to suggest that despite the relation between words and ideas, it always remains possible – and generally even desirable – to reason independent of linguistic convention. Yet as Condillac presents words and ideas as necessarily connected, reason becomes almost inextricable from language and its conventions.

Combining Hobbes's position on the relation between words and ideas with Locke's reflections on the status of language as a tacitly upheld social convention, Condillac thereby arrives at an explicit formulation of linguistic relativism. As language is a communal institution upheld by tacit consent, ideas in the minds of individuals tend to reflect the conventions of their language.[318] Through the medium of language, the mind therefore acquires a conventional structure and set of terms with annexed concepts that are largely shared with other members of the speech-community. Condillac thereby presents the mind as a socialised entity shaped by the conventions of the language it has learned. Yet Condillac maintains that the linguistic conventions may not only determine the composition of individual ideas, but also that the structure of language may influence our train of thought in other ways. Thus Condillac claims that due to their overall structure, some languages are more suitable for certain forms of thought and expression than others. For instance, Condillac maintains that ancient languages more perfectly facilitated poetic expression, while modern

[316] Condillac, *Essay* II.i.1 §1-2 p.114-115; I.iv.2 §17-20 p.85-87.

[317] Cf. Martine Pécharman, 'Signification et langage dans l'Essai de Condillac', *Revue de Métaphysique et de Morale* 1 (1999) 81-103, 96-99.

[318] Condillac, *Essay* I.iv.2 §25 p.90: 'Since human beings cannot make signs for themselves except when living together, it follows that the fund of their ideas, when their mind begins to be formed, consists entirely in their mutual intercourse.'

languages like French are more suitable for scientific investigation and communication.[319] This last theory is then lends additional weight to the conclusion that Condillac thought that the conventional structure of languages has the potential of determining or influencing the thought of individuals.[320]

The History and Imperfections of Language

The inextricable connection between language and thought in Condillac's theory is particularly significant in light of his conception of language as a historical and fluid construction. Condillac thereby develops another of Locke's suggestions, namely that language is the product of a long and somewhat haphazard historical development. Thus we have seen that Locke worries that linguistic conventions have been created in a bygone era by people largely indifferent or ignorant of scientific or philosophical concerns. For this reason, Locke notes that commonplace linguistic conventions are often unsuitable for use in scientific or philosophical discussions.[321] Yet while Locke had briefly touched upon this subject, the historicity of language becomes a major preoccupation for Condillac, as he constructs a conjectural history of the origin and early development of language. Condillac maintains that language first arose out of the spontaneous cries and gestures of natural man. These first signs constituted what Condillac calls 'the language of action' and served to communicate natural man's primary needs and sentiments.[322] Building upon the language of action, primitive man gradually instituted conventional signs to denote experiences and purely intellectual concepts in ever greater detail – thereby slowly developing the language of civilised man.[323]

[319] Condillac, *Essay* II.i.15 §147-156 p.187-192.

[320] Charles Taylor, *The Language Animal: the Full Shape of the Human Linguistic Capacity* (Cambridge MA 2016) 18-19 disputes this analysis and argues that Condillac did not truly arrive at a position of linguistic relativism. Taylor maintains that like Locke, Condillac is solely concerned with the relationship between signs and ideas, but not with the connections between words among each other or with the structure of language. As we have seen, this interpretation overlooks the fact that Condillac clearly does maintain that signs connect among themselves and function as a means of connecting ideas. Nonetheless, Taylor is correct that Condillac does not develop this perspective as far as later thinkers like Herder or Hamann, who unlike Condillac deny that there is really a way of escaping the scope of language.

[321] Locke, *Essay* III.vi.25 p.452-453.

[322] Condillac, *Grammaire* I.i in: *Oeuvres Philosophiques* ed. by Georges le Roy (Paris: Presses Universitaires de France 1947-1951) I p.428-429.

[323] Condillac, *Essay* II.i.1 §3-9 p.115-117. Cf. Condillac, *Logic* II.ii p.390. Rousseau was not fully convinced by his friend's conjectural history. Although Condillac underscored that the development

Thus, Condillac – again echoing Locke – stresses the fact that language was first developed by men largely unconcerned with its accuracy. The natural cries and gestures of the language of action had been adequate representations of the sentiments they signified. But when primitive man, incited by his passions, ventured to construct signs not directly related to securing his survival:

> Then men felt less and less the need to analyse. Soon, they felt only the desire to speak, and they spoke before having any idea what they wanted to say. Judgment was no longer naturally put to the test of experience. They did not have the same stake in making sure that the things they judged were as they had supposed. They wanted to believe them without examination. Judgments that had become habitual become opinions that were not questioned. [...] Languages thus became very defective methods.[324]

According to Condillac, language thus has been developed carelessly, without taking into account all the facts of experience. He singles out figurative and allegorical expressions as especially common sources of error, as they are often taken literally once their original meaning is forgotten.[325] Nonetheless, Condillac postulates that over the course of recent history, languages may have become more precise due to the general progress of enlightenment. Incidentally, this is also the reason why modern languages are more suitable for scientific discourse,

---

of language was a gradual and haphazard process, Rousseau inquired how natural man could have possessed the intellectual abilities to develop language if the required reflection and reasoning depend on language itself. On the face of it, Rousseau makes a fair point – although probably this boot-strapping problem is ultimately more damaging to the hypothesis that language and thought are inextricably linked than to Condillac's rendering of the history of language. Yet it is not obvious that Condillac considered language – or at least language as it had developed naturally – as the product of careful reflection. Instead, he emphasised that language was the product of a long evolution, in which chance and contingency had contributed as much as conscious design. See Rousseau, *Second Discourse* in: *The* Discourses *and other early political writings* ed. and trans. by Victor Gourevitch (Cambridge: Cambirdge University Press 1997) 145-149; Cf. Hans Aarsleff, 'Pufendorf and Condillac on Law and Language', *Journal on the Philosophy of History* 5 (2011) 308-321, 318-319 and Avi Lifschitz, *Language and Enlightenment: The Berlin Debates of the Eighteenth Century* (Oxford 2012) 26-29 for a discussion of Condillac's account of the evolution of language and Rousseau's critique.

[324] Condillac, *Logic* II.iii p.393.

[325] Condillac, *Essay* II.i.10§104 p.166. Cf. McNiven Hine, *A Critical Study of Condillac's* Traité des systemès 192.

while ancient languages, full of metaphors and figurative expressions, lend themselves more perfectly to poetic expression.[326]

Yet despite these advances, many of the imperfections accumulated throughout history prevail in modern languages. What is more, Condillac deplores the fact that even today, linguistic conventions are continuously revised, more often making language less rather than more precise:

> Mais on ne déterminera pas la signification des mots : on l'altérera, on la changera sans raison : une métaphore, une comparaison paroîtra répandre la lumière ; et, pour expliquer une expression qu'on n'entendra pas, on en imaginera d'autres qu'on n'entendra pas davantage. C'est de la sorte que d'un langage confus naissent des opinions ; et que de ces opinions naît un autre langage, qui, tout aussi confus que le premier, produit de nouvelles opinions, pour produire bientôt de nouveaux langages également confus : et ainsi de suite, pendant des siècles.[327]

This passage displays Condillac at his most pessimistic. Human beings may have created the extraordinarily useful institution of language as a method for both communication and the extension of thought. Yet apparently, a majority of people cannot be bothered to use this method correctly. Not merely disregarding established usage, they are often completely unconcerned with establishing representative and precise linguistic conventions. Over time, language has therefore become a very inaccurate – even misleading – representation of empirical reality.[328] This situation is especially distressing in view of the close connection between language and thought in Condillac's epistemology. It invites the kind of question Rousseau would later pose: whether the deceptions and inconsistencies of the conventional language of civilised man might not ultimately derogate its usefulness.

---

[326] Condillac, *Essay* II.i.15§156 p.191-192. Cf. James Stam, 'Condillac's Epistemolinguistic Question', in: R.W. Rieber (ed.), *Psychology of Language and Thought* (New York: Plenum Press 1980) 77-90, 84-85.

[327] Condillac, *Histoire Ancienne* III.xxvii in: *Oeuvres Philosophiques* ed. by Georges le Roy (Paris 1947-1951) II p.91. In this thesis, we will read French primary sources in English when a translation of sufficient academic standard is available. As this is not the case for all French texts cited, we will render some quotations in the original language.

[328] Condillac, *Essay* I.iv.1 §10 p. 82: 'Anyone who looks into himself will find a large number of signs which he has only connected with very imperfect ideas, and even some to which he has connected none at all. That is the source of the chaos that prevails in the abstract sciences, a chaos the philosophers have never been able to clear up because none of them has known its first cause.'

The Method of Analysis

Yet Condillac himself is not as pessimistic as his friend from Geneva, as he thinks that both individuals and society as a whole are be able to either circumvent or rectify the errors and imprecisions of language. Thus Condillac claims that individuals would be able to avoid the deception of words by employing what he refers to as the method of analysis. Noting that linguistic conventions are often misguided, Condillac first of all maintains that simply consulting the conventional definitions of terms is not a reliable method for determining the composition of our ideas.[329] Alternatively, Condillac maintains that in order to reason well, an individual must correlate their ideas to experience rather than rely upon words and take their conventional or assumed signification at face value:

> Consequently, the only means of acquiring knowledge is to return to the origin of our ideas, follow their generation, and compare them in terms of all their possible relations. That is what I call to "analyse."[330]

Before employing our ideas in reasoning, we should thus decompose them into their constitutive parts to find out their precise structure and mutual relations.[331] Subsequently, Condillac states that we should 'return to the origin of our ideas.' As Condillac believes that our sensations of empirical reality are ultimately the only reliable source of ideas, we should compose or recompose our ideas with the help of experience. Besides advocating the scrutiny of received opinion and the conventional signification of words, Condillac's method of analysis and directive to 'return to the origin of ideas' thereby entails the rigorous empirical investigation of reality.[332]

Condillac contrasts his own method of analysis with the methodology employed by many seventeenth century philosophers including Descartes, Malebranche, Spinoza and Leibniz. According to Condillac, these thinkers used the method of synthetic reasoning rather than analysis to construct elaborate

---

[329] Condillac, *Logic* II.i p.386; *Essay* II.ii.2 §19 p.203; II.ii.2 §11 p.200; Cf. *Essay* I.iii. §11 p.74 for Condillac's reservations about definitions.

[330] Condillac, *Essay* I.ii.7 §67 p.49; Cf. II.ii.3 §39 p.213-214.

[331] Condillac, *Grammaire* I.iii *OPh* I p.435-436.

[332] Cf. André Joly, 'De la théorie du langage à l'analyse d'une langue', in: Sgard (ed.), *Condillac et les problèmes du langage* 243-256 for a more elaborate discussion of Condillac's analytical method.

philosophical systems. Following the methodology of geometry, they predicated their reasoning on a collection of fundamental axioms that, in Condillac's view, are often assumed rather distilled from experience.[333] In addition, Condillac criticises seventeenth century philosophers for using terms denoting abstract ideas – such as 'substance' or 'essence' – without sufficiently determining and specifying their meaning with the help of experience. In fact, following his rejection of innate ideas, Condillac accuses his illustrious predecessors of simply inventing these abstract ideas fundamental to their philosophy, describing them as products of the imagination rather than investigation.[334] Providing a detailed refutation of the philosophy of Spinoza, Leibniz and Descartes in his *Treatise on Systems*, Condillac concludes that due to their adoption of synthetic method and reliance on unsubstantiated abstract ideas, these thinkers created an intricate philosophical edifice that upon further inspection turns out to have no relation to empirical reality.[335]

According to Condillac, his analytical approach will preclude the mistake made by both these seventeenth century philosophers and primitive man developing the conventions of language – the unprincipled and unsubstantiated composition of ideas. Our foremost priority is therefore the scrutiny of the ideas and relations we attach to signs and employ in our thoughts. Only after we have adequately recomposed our ideas in reference to experience, are we in a position to reason accurately by perceiving the relations between ideas. Staying as close to experience as possible, we will not be tempted in our scientific and philosophical investigation to adopt or devise fanciful notions on topics beyond the reach human understanding.[336] Condillac thereby advocates proceeding from the

[333] Condillac, *Essay* I.ii.7 §61-64 p.46-47; Condillac, *Treatise on Systems* XII p.123. Oddly enough, Isabel Knight has entitled her admirable but somewhat dated book on Condillac's philosophy *The Geometric Spirit: The Abbé de Condillac and the French Enlightenment* (New Haven 1968), despite the fact that Condillac was fundamentally opposed to the methodology of both seventeenth century philosophers modelling their approach to that of geometry, as well as to that of geometry itself.

[334] Cf. T. Takesada, 'Imagination et Langage dans l'Essai sur l'origine des connaissances humaines de Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 47-58, 52-53.

[335] For a critical discussion of Condillac's refutation of seventeenth century philosophers, Cf. McNiven Hine, *A Critical Study of Condillac's* Traité des systèmes 48-121.

[336] Condillac, *Essay* II.ii.3 §30 p.209: 'By reasoning according to this method, only with simple ideas or with complex ideas that are the work of the mind, we will have two advantages: the first is that since we know the generation of the ideas on which we meditate, we will not move forward without knowing where we are, how we got there, and how we can retrace our steps. The second is that on every subject we will clearly see the limits of our knowledge, for we will be up against them

particular facts of reality towards general and abstract ideas, rather than to perceive and explain empirical phenomena with the help of preconceived abstract notions, like many seventeenth-century philosophers had supposedly done.[337] Condillac's method of analysis thus turns out to be a form of empiricism not too dissimilar from Locke's epistemology. As Locke before him, Condillac concludes that individuals should not rely on the conventional signification of words, but rather remain as close to experience as possible in their reasoning.

However, it is not clear if Condillac's method of analysis complies with his position that thought and language are interdependent. As we have seen, Condillac maintains that complex ideas cannot be conceived without the use of linguistic signs. Therefore any form of thought that exceeds the level of mere instinct requires the use of language. As the method of analysis described by Condillac clearly qualifies as a higher form of cognition, the process of consciously deriving ideas from experience while ignoring or rectifying the meaning of words must itself rely on language and its cognitive function. This raises an issue familiar to many philosophers arguing in favour of linguistic relativism. If we do not possess any autonomous or innate faculty of the mind, and our thought is indeed enabled and shaped by language, how would we be able to escape its influence? Following Condillac's method of analysis, we may try to ignore the conventions of language and formulate our ideas in reference to experience. Yet due to Condillac's disavowal of Locke's tacit assumption that ideas are composed with the help if an innate faculty of the mind, it is doubtful whether his epistemology still allows for this approach of thinking outside linguistic conventions. For in the absence of an innate cognitive faculty, the method of analysis itself would have to depend on complex ideas integrated with terminology with a conventional signification.

The Analysis of Moral Concepts

In our previous chapter, we have seen that while Locke advocates composing ideas of substances by means of experience, he does not present a clear method for the composition of moral concepts. As Condillac's method of analysis is

---

when the senses cease to furnish us with ideas so that the mind consequently can form no further notions.'

[337] Condillac, *Essay* I.v §5 p.93; Condillac, *Treatise on Systems* I p.3; *Treatise on Sensations* Introduction p.169.

modelled on Locke's epistemology, this raises the question whether Condillac in fact does provide a consistent approach for the formation of moral notions. To begin with, the fundamentals of Condillac's moral epistemology are again borrowed from Locke. Thus Condillac presents moral judgment as the outcome of reasoning based on moral concepts. Condillac also retains Locke's category of ideas called mixed modes – ideas made freely by the mind without regard to the objects of empirical reality.[338] Furthermore, Condillac restates Locke's contention that as the mind itself composes these ideas of modes, it may perceive all possible relations between ideas and reason with absolute certainty on moral matters.[339] Finally, Condillac maintains that like any other kind of idea, moral concepts depend on linguistic signs for structural stability and the establishment of mutual relations between ideas. In fact, with moral concepts, this dependency on words is even stronger, as mixed modes do not depend on any examples in nature, making them especially reliant on linguistic signs.[340] Accordingly, Condillac envisions that the conventional signification of moral terms appears as the most immediate standard from which individuals often derive their moral concepts.[341]

Yet as in the case of ideas of substances, Condillac thinks that the standard of common use is often an unsuitable guideline for the composition of moral concepts. In fact, Condillac maintains that the greater dependency of morality on language has caused a particularly serious confusion in our moral notions.[342] Accordingly, it seems that of all ideas, mixed modes are particularly in need of scrutiny through the method of analysis. Yet it is not immediately clear

[338] Condillac, *Essay* I.iii. §5 p.72.

[339] Condillac, *Essay* I.iii. §15 p.76. In the same passage, Condillac also repeats Locke's claim that as the same cannot be said about complex ideas of substances, our reasoning about empirical objects cannot claim the same degree of certainty. But whereas this position had lead Locke to a qualified scepticism regarding knowledge of empirical reality, Condillac does not stress the issue and generally appears to be much less concerned with any sceptical consequences. Cf. Condillac, *Essay* I.v §12 p.99 for an example.

[340] Condillac, *Essay* I.iv.8 p.81: 'If you think you do not need words, pull them out of your memory and try to reflect on civil and moral laws, on virtues and vices, in short on all human actions, and you will see how mistaken you were.' Cf. II.i.1 §106 p.167.

[341] Condillac, *Essay* II.i.1 §106 p.167. Cf. II.i.15 §162 p.194.

[342] Condillac, *Essay* II.ii.2 §19 p.203: 'What on these occasions raises numerous obstacles to fixing the sense of terms, even after many efforts leaving much equivocation and obscurity behind, is that we take words as we find them in the usage to which we absolutely intend to conform. Moral philosophy especially is full of intricately composed expressions, and the usage we consult is so inconsistent that this method will unavoidably make us speak with little exactness and cause us to fall into a great many contradictions.' Cf. II.ii.2 §11 p.200.

how Condillac thinks ideas of morality are to be analysed and reconstructed. He reiterates that individuals should first examine precisely the ideas they attach to certain terms and be ready to specify their usage to their interlocutors.[343] Condillac's method of analysis subsequently requires individuals to scrutinise the composition of their ideas and revise them if necessary. But as mixed modes do not depend on models in nature, the question arises in what manner they may be rectified. Condillac's offers a somewhat enigmatic answer, as he suggests that our moral concepts may nonetheless be refined through experience:

> With these details about mixed modes, it is easy to understand that it is entirely up to us to fix the meaning of their names, because it is our task to decide on the simple ideas which we have ourselves put together in the collections. We also understand that other people will share our thoughts, provided they are placed in the circumstances in which the same simple ideas are the object of their minds as of ours, and in which they are induced to unite them under the same names under which we have collected them.[344]

Condillac seems to suggest that when placed in similar circumstances, individuals are exposed to the same sensations, and consequently will be induced to construct their moral concepts in comparable ways. Despite the fact that these moral concepts are not based on concrete models in nature, they are nonetheless somehow based on experience. Far from clarifying the process for the analysis of moral ideas, this passage raises the question how experience alone could constitute the basis for the reliable formulation of moral concepts.

To answer this question, we should briefly reconsider Condillac's account of the properties of experience. While Condillac maintains that by providing sense-data about the objects of empirical reality, experience is the sole reliable source of ideas, he also claims our sensations are necessarily intermixed with feelings of pleasure or pain. As we have seen, Condillac subsequently claims that these feelings of pleasure and pain provide the underlying motivation both for our actions as well as for the development of knowledge and our intellectual faculties. In addition, Condillac maintains that the fulfilment of the desire for pleasurable sentiments is our goal in life, while happiness consists of the achievement of this goal. Thus Condillac states that: 'on est heureux, toutes les

[343] Condillac, *Essay* II.i.11 §116 p.171.
[344] Condillac, *Essay* II.ii.2 §26 p.207.

fois qu'on chasse un besoin par des sentimens agréables.'[345] Following Hobbes and Locke, Condillac thereby unmistakably retains a hedonist conception of the good. Accordingly, the ultimately the goal of moral guidelines from the perspective of the individual would be to steer us away from pain and achieve pleasure. For this reason, some commentators have in fact classified Condillac as an early proponent of utilitarianism, ascribing to him the view that the purpose of morality is simply the maximisation of pleasure among individuals.[346]

Yet a closer examination of Condillac's moral thinking reveals that his position that individuals are necessarily motivated by the passions does not necessarily lead him directly to the conclusion that the purpose of morality and society is the greatest happiness of the greatest number. Rather, following Hobbes's example, Condillac presents the fact that individuals are motivated by passions as an argument for the necessity of instituting morality and centralised authority. Thus Condillac points out that our goal to maximise pleasure and avoid pain is most effectively obtained – at least in the long run – by acting sociable and taking the pleasures and pains of others into account:

> L'expérience ne permet pas aux hommes d'ignorer combien ils se nuiraient, si chacun, voulant s'occuper de son bonheur aux dépens de celui des autres, pensait que toute action est suffisamment bonne dès qu'elle procure un bien physique à celui qui agit. Plus ils réfléchissent sur leurs besoins, sur leurs plaisirs, sur leurs peines, et sur toutes les circonstances par où ils passent, plus ils sentent combien il leur est nécessaire de se donner des secours mutuels. Ils s'engagent donc réciproquement ; ils conviennent de ce qui sera permis ou défendu, et leurs conventions sont autant de lois auxquelles les actions doivent être subordonnées ; c'est là que commence la moralité.[347]

Condillac's theory that humans are motivated by the desire for pleasure and aversion of pain thus leads him to the conclusion that man most optimally achieves the satisfaction of his needs and appetites by living together in society

[345] Condillac, *Histoire Ancienne* III.xxiv *OPh* II p.81.

[346] The argument that Condillac is an early proponent of utilitarianism is made by Arnoud Orain, 'The Moral Theory of Condillac: A Path toward Utilitarianism', *Revue de philosophie économique* 13 (2012) 93-117 and Ulrich Ricken, 'Linguistique et anthropologie chez Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 75-93, 77. Yet whether one would agree with this analysis is largely dependent on the definition of utilitarianism one would choose to adopt.

[347] Condillac, *Traité des animaux* II.vii p.180.

and providing mutual assistance.[348] This conclusion is likely inspired by the conception of sociability proposed by Samuel Pufenderof, which was particularly influential in eighteenth century France.[349] Pufendorf maintained that a rational consideration of our needs, faculties, and circumstances would lead humans to conclude that providing mutual assistance and living in society is in their collective best interest. Rejecting the notion that humans are a social beings by nature, Pufendorf rather claimed that man's sociability is the consequence of prudential reasoning.[350]

Even if Condillac appears to borrow his account of sociability from Pufendorf, he diverges from his predecessor in his description of the method by which humans learn to become sociable. Whereas Pufendorf, as well as Hobbes and Locke before him, had all claimed that natural law is accessible to human reason, Condillac states in the passage quoted above that it is *experience* that teaches humans that their well-being depends on following a set of rules that allows them to live together.[351] But what does this entail? As we have seen, experience is for Condillac not limited to the perception of empirical objects, as feelings of pleasure and pain are always intermixed with sensations. As the objective of morality is to guide us in acquiring or avoiding precisely these feelings, it follows that to some extent, moral precepts may be learned by trial and error – through experience:

> Les conditions, que ces conventions renferment, sont les premières lois des sociétés. On les peut nommer *lois naturelles*, parce que l'homme n'a pas besoin de méditer pour les découvrir. Tout lui apprend qu'il ne doit pas nuire, s'il ne veut pas qu'on lui

[348] Condillac, *Histoire Ancienne* VI p. 42: 'Trop foibles pour veiller séparément à leur conservation, les hommes ont été forcés par les circonstances à se donner mutuellement des secours. Plusieurs ont donc consenti à vivre ensemble, et cet accord est le premier fondement des sociétés. La fin qu'ils se proposent, est que leur union soit avantageuse à chacun en particulier et à tous ensemble ; c'est à cette condition qu'ils s'unissent. Il s'agit donc pour eux de concilier les intérêts différens, et de les faire concourir à un seul et même intérêt général.' Cf. *Histoire Ancienne* III.xxv *OPh* II p.83; I.v *OPh* II p.15.

[349] On Pufendorf's influence, see Schneewind, 'Pufendorf's Place in the History of Ethics', and Pagden, *The Enlightenment* 56-63 for a discussion of Pufendorf's importance to the *philosophes*.

[350] Pufendorf, *On the Duty of Man and Citizen* ed. by James Tully and trans. by Michael Silverthorne (Cambridge 1991) I.iii.4-8 p.33-35.

[351] Pufendorf, *On the Duty of Man and Citizen* p.17-18; I.iii.12 p.37; Locke, *Two Treatises of Government* II.ii.12 p.105.

> nuise : et qu'il doit secourir s'il veut être secouru. L'expérience suffit pour lui enseigner ces maximes, et elle les lui confirme tous les jours.[352]

Of course, Condillac affirms that the experience of pleasure and pain should be subject to reflection, enabling a more methodical and comprehensive formulation of moral precepts. Nonetheless, Condillac maintained that the experience of pleasure and pain provides an elementary indication of morality and natural law. Accordingly, Condillac states that 'pleasure and pain are our first teachers. They enlighten us because they inform us whether we are judging badly or well.'[353]

Condillac's theory that a conception of morality may be derived from the experience of pleasure and pain ultimately remains somewhat equivocal. Condillac seems to suggest that humans simply acquire a conception of morality from an accumulated experience of pleasure and pain. Yet this theory leaves several ethical questions largely unanswered. Thus if individuals acquire a complete understanding of morality merely by means of sensation, why do humans often tend to diverge from the prudential guidelines experience has provided them with? More importantly, as Condillac thinks that morality is derived from personal experience, he presents morality largely as an individualistic prudential guideline. This raises the question how morality would be able to regulate the behaviour of individuals living together in society. Thus given the variety of experience among individuals, how could their presumably diverging feelings of pleasure and pain give rise to a moral code common to all members of society? Unfortunately, Condillac leaves these questions largely unanswered.

The Reform of Education

In the preceding, we have seen that Condillac thinks that the ideas of both substances and of morality in the minds of individuals are at least to some extent determined by the conventions of language. We have also seen that Condillac presented his method of analysis as an alternative means to acquire ideas, thereby circumventing the generally imprecise or inaccurate standards of common use. Yet there is some indication that Condillac realised that by itself the method of analysis is not sufficient to circumvent the pernicious influence of language. Thus Condillac presents two proposals to minimise – or even eradicate

[352] Condillac, *Histoire Ancienne* I.vi *OPh* II p.16.
[353] Condillac, *Logic* I.i p.349.

– the potential of language to deceive and misinform individuals and society as a whole. Firstly, Condillac maintains that the advancement of knowledge requires the reform of education. Condillac's preoccupation with education is evidenced above all by his *Cours d'études*, a collection of study books in thirteen volumes, written in the years 1767-1773 for the young Ferdinand of Parma, grandson to Louis XV.[354] Of course, many early-modern empiricist philosophers including Locke, Helvétius and Rousseau were also preoccupied with education – and perhaps not surprisingly. The position that all our ideas are acquired externally – either through experience or linguistic communication – almost immediately raises a concern for exposing individuals to the right kind of ideas.

In his philosophical writings, Condillac had already discussed this same issue. As an empiricist, Condillac denies that human beings possess an innate faculty of reason that would allow them to perceive eternal truths independent from experience. Consequently, the scope and outcome of our reasoning is constrained both by our intellectual habits and the ideas we retain or acquire. As Condillac thinks that many of our ideas are learned from others in childhood, we are prone to make the same mistakes as our teachers:

> Because in childhood, we get out thoughts from others, we adopt all their prejudices. When we reach an age where we believe we think for ourselves, we still continue thinking second hand, because we think according to the prejudices they have bequeathed us. Under these conditions, the more progress the mind seems to make, the further afield it strays, and errors pile up from generation to generation. When matters reach this point, there is only one means of restoring order to the faculty of thought – to forget everything that we have learned, to take up our ideas again at their origin, and to follow their elaboration and to reconstitute, as Bacon put it, human understanding.[355]

Furthermore, as prejudices are not only transferred through language, but actually inhere in the very linguistic conventions themselves, Condillac concludes that instruction through speech and writing may ultimately do us more harm than good. As children have not yet developed sufficient capacity for analysis – which in itself requires the use of language – they will not be able to scrutinise

[354] Cf. Louis Trénard, 'L'influence de Condillac sur l'enseignement', in: Sgard (ed.), *Condillac et les problèmes du langage* 145-164 for a discussion of Condillac's influence on French educational thought, focusing on his impact in the late eighteenth and early nineteenth century.

[355] Condillac, *Logic* II.i p.386.

and correct the notions acquired through instruction.[356] Condillac claims that many of these notions – and the linguistic signs that transfer them – accommodate error and prejudice. Through education, these prejudices become ingrained to the point that they are often conceived as self-evident truths or innate ideas.[357]

Yet Condillac is not merely apprehensive about the prevalence of prejudice and error in both the curriculum and in language. Following his position that cognitive faculties are not innate but gradually acquired, Condillac is also concerned that current educational practice inculcates the wrong intellectual habits, noting that 'our mind becomes used to this lack of care about correctness by the way we grow into language.'[358] Thus Condillac claims that initially, infants acquire their ideas through experience. It is only when learning the use of language that children become used to receiving ideas through the inaccurate and confusing conventions of signification, subsequently contracting the habit of thinking upon words rather than ideas.[359] According to Condillac, it is therefore crucial that pupils are taught to reason according to the guidelines of his method of analysis – by scrutinising terms and acquiring ideas from experience. As he wrote a two-part instruction manual as part of his *Cours d'études* on the proper conduct of the understanding entitled *L'art de penser* and *L'art de raisonner*, Condillac believed that older pupils, who had acquired sufficient reflection, are to be instructed in the use of the method of analysis. Younger children however, should not be taught any subjects that are beyond the comprehension of their age. Prefiguring a theme further developed by Rousseau in his *Émile*, Condillac therefore maintains children should be encouraged to acquire ideas not through formal instruction, but rather by experience, through interacting playfully with people and the objects of empirical reality. It is only by this method that they acquire the intellectual habit of referring to experience rather than language in their reasoning.[360]

[356] Condillac, *Essay* II.ii.3 §29 p.208-209.
[357] Condillac, *Essay* I.ii.9 §79-80 p.55-56; II.ii.1 §3-6 p.197-198; Condillac, *Logic* I.i p.349.
[358] Condillac, *Essay*; II.ii.1 §5 p.197.
[359] Condillac, *Logic* I.i p.348.
[360] Condillac, *Essay* I.ii.5 §50 p.42; II.ii.3 §42 p.215; I.iv.1 §10-12 p.82-83; *Logic* I.i p.348.

The Perfection of Language

While the reform of education may be able to reduce the spread of the most obvious prejudices, it leaves the conventions of language itself largely intact. As we have seen, Condillac maintains that these conventions are often inaccurate and confused representations of empirical reality. With his method of analysis, Condillac presents a method by which individuals can scrutinise and – if necessary – correct their ideas independent from the conventions of language. Yet not everyone is equally willing or able to apply analysis consistently to each and every idea acquired. Condillac therefore maintains that the most effective means to ensure the dissemination of accurate knowledge would be to reform language itself:

> I have observed that to attain true knowledge we should start afresh in the sciences without allowing ourselves to become biased in favour of trusted opinions, and for this reason it has seemed to me that to make the language exact, we must reform it without regard to usage.[361]

Due to its constitutive role in the formation of the mind and its ideas, the accuracy of our knowledge becomes inextricably linked to the constitution of our language. According to Condillac, any attempt to improve or enlarge scientific knowledge or philosophical insight, should therefore be accompanied by a rigorous reform of our linguistic conventions.

Condillac's declaration that language should be reformed was not merely an idle claim. Later in life, he developed his proposal into a concrete program for the perfection of language. As part of his *Cours d'études* for the Prince of Parma, Condillac not only wrote a comprehensive *Dictionnaire des synonymes*, he also presented a manual for the analysis and reform of language in his *Grammaire*. The main object of this reform is to fashion language into an accurate and ordered – rather than confused and misleading – signification of our ideas:

> Les langues ne se perfectionnent qu'autant qu'elles analysent ; au lieu d'offrir à la fois des masses confuses, elles présentent les idées successivement, elles les distribuent avec ordre, elles en font différentes classes ; elles manient, pour ainsi dire, les élémens de la pensée, et elles les combinent d'une infinité de manières ; c'est à quoi elles réussissent plus ou moins, suivant qu'elles ont des moyens plus ou moins

[361] Condillac, *Essay* II.ii.2 §11 p.200; Cf. *Essay* II.ii.4 §53 p.220.

> commodes pour séparer les idées, pour les rapprocher, et pour les comparer sous tous les rapports possibles.[362]

Condillac's ultimate objective is thus to reform language to reflect the results of an analysis of our ideas. After examining our ideas and mutual relations, we should revise the conventions of language to conform to the organisation of the ideas in our mind. The result will be a well-formed language, in which signs will precisely signify ideas and clearly indicate their relations. Consequently, language *itself* will become a reliable method of analysis. Condillac presents algebra as an example of such a well-formed language, as each sign exactly represents a well-defined idea and the connections between its signs are immediately apparent, thereby also displaying the relations between associated ideas. In his *Langue des calculs*, unfinished at his death and published posthumously only in 1798, Condillac finally attempted to put his proposal into practice, by designing a flawless language modelled on algebra.[363]

This comparison of his perfect language with algebra has led some interpreters to claim that later in life Condillac diverged from his empiricism and intended to reform language in line with universal logical principles.[364] Thus, John O'Neal claims that in view of Condillac's preoccupation with the perfection of language – which he shared with Leibniz and Bacon among others – we would be justified to classify him as a 'rationalist' thinker.[365] Yet a number of other commentators have rightly questioned both this latter characterisation and the notion that later in his career, Condillac departed from his empiricist principles to become more interested in formal logic.[366] They point out that Condillac may have had the goal of devising a universal language that would encapsulate universal logical rules, thereby perfecting the function of human understanding. But these logical rules are not derived from a set of abstract principles or criteria. Rather, these rules should be deduced from experience – they are the

[362] Condillac, *Grammaire* I.iii *OPh* I p.435

[363] Condillac, *Logic* II.vii p.410; *La Langue des Calculs* II.i *Œuvres Philosophiques* II p.471. Cf. Alain-Marc Rieu, 'Le complexe nature – science – langage chez Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 27-46, 38; Duchesneau, 'Condillac et le principe de liaison des idées' 62.

[364] Wojciechowska, 'Le sensualisme de Condillac', 318;

[365] John O'Neal, *The Authority of Experience: Sensationist Theory in the French Enlightenment* (University Park 1996) 35, 59. Unfortunately, O'Neal declines to explain what he means precisely with this rather broad characterisation – other than by mentioning that Condillac's method for the perfection of language apparently has similarities to proposals made by Leibniz.

[366] Nicholas Rousseau, *Connaissance et langage chez Condillac* (Geneva 1986) 28.

consequence of exterior phenomena well analysed.[367] Ultimately, Condillac's goal was therefore to create a language that would precisely and flawlessly reflect and represent the ideas in the mind of speakers, which in turn should be derived from the facts of experience. His proposal for reform is to be carried out by subjecting both language and our ideas to analysis – the empiricist method Condillac himself had laid out. Even Condillac's later attempt to design a perfect language based on the example of algebra was therefore perfectly in line with his position that all ideas should be scrutinised with the help of experience.

Yet in the end, Condillac's plan for the perfection of language raises the same question as his method of analysis. We have seen that according to Condillac, language plays a defining role in the structuring ideas and development of our cognitive faculties, thereby giving rise to the ability for reasoning. Yet the question is how the scrutiny and reform of language could be undertaken if not with the help of rational reflection, which according to Condillac's theory, again depends on language – precisely the object to be reformed. To be sure, Condillac maintains that this reform of language should be conducted in reference to experience, rather than through purely reflective reasoning. Yet even the interpretation of experience requires a certain analytical method. And as we have seen, Condillac claims that language itself functions as this analytical method. In order to reform or scrutinise language on the basis of experience, one would therefore require the ability of thinking independent of the conventions of the language one is trying to reform. This kind of purely mental, non-linguistic reasoning is still accommodated in Locke's theory of the understanding. But the whole point of Condillac's philosophy had been to show that reason is not an innate faculty; that it is a habit developed out of sensation and depends on language as an analytical method. Condillac thus ascribes to language such an integral role in the establishment of reasoning that it is difficult to conceive how his objective of scrutinising and perfecting language could be achieved, as it turns out that linguistic conventions themselves are a necessary prerequisite to the procedure of their own reform.

Ultimately, Condillac's insistence that the mind has no innate cognitive faculties, but that its abilities are merely habits developed through sensation and with the help of language, creates a fundamental tension within his philosophy.

[367] Rieu, 'Le complexe nature – science – langage chez Condillac', 35-37; Duchesneau, 'Condillac et le principe de liaison des idées', 56, 63; Sylvain Auroux, 'Empirisme et théorie linguistique chez Condillac', in : Sgard (ed.), *Condillac et les problèmes du langage* 177-219, 182-183.

On the one hand, this position allows Condillac's philosophy to be interpreted as leading to a form of linguistic determinism according to which all human actions are occasioned by the interaction between sensations and accumulated experience organised through language. Yet on the other hand, the fact that Condillac urges us to transcend and remedy the conventions of language means that he somehow still assumes that humans possess a reflective faculty that functions independently or beyond the collection of ideas structured through language. Yet as he claims that our reflective faculty actually comprises nothing more than ideas acquired through sensation structured through language, it remains unclear how Condillac is able to account for such a cognitive faculty in his philosophy.

Conclusion

In the preceding, we have seen that while Condillac derives many of his theories from Locke, he nonetheless revises the philosophy of his illustrious predecessor on at least two crucial points. On the one hand, Condillac dispenses with Locke's tacit assumption that the mind is an autonomous entity conducting mental operations by apprehending or manipulating its ideas. Presenting the mind as devoid of any innate faculties, Condillac rather maintains that all forms of thought are ultimately nothing but cognitive habits established by means of experience. On the other hand, Condillac adopts Locke's suggestion that words have an important function in enabling reason, extending it to the point that language becomes responsible for all higher forms of cognition. Subsequently, we have seen that the theory that language enables human cognition then leads Condillac towards the conclusion that the conventions of our language may in fact influence or determine our thoughts. Yet we have discovered that by simultaneously eradicating all innate cognitive faculties and presenting language as the primary explanation for human reason, Condillac has introduced a fundamental tension into his philosophy. We have seen that Condillac maintains that as modern language has been disfigured by prejudice, linguistic conventions provide defective standards for reasoning. Yet as Condillac claims that human cognition is entirely dependent on language, it is unclear whether his philosophy would in fact allow humans to reason independent from these conventions of language.

Despite this tension, Condillac's philosophy does provide novel ideas and perspectives that would influence a number of subsequent thinkers. For instance,

we will see in our next chapter that Helvétius adopted Condillac's conception of human understanding as devoid of any innate faculties and concluded that the mind is nothing than the collection of its ideas. Condillac's philosophy is therefore an important step in the development of empiricism from the Lockean position that all ideas are derived from experience towards the reductionism of *philosophes* like Helvétius and d'Holbach, who claim that all human cognition is nothing more than sensation in different forms. Furthermore, in our final chapter we will see that Condillac's philosophy also had a profound influence on Rousseau, who adopts the theory that man's cognitive faculties and moral outlook are the combined product of experience and the use of signs and concludes that like language, human nature must also have had a history. For this reason, it would not be an exaggeration to maintain that without Condillac's inspiration, Rousseau would not have conceived his conjectural history of the development of human nature as outlined in the *Second Discourse*.

4. Empiricism at its Limits: Helvétius's Reductionist Conception of Human Nature

Introduction

Claude Adrien Helvétius develops his philosophy on empiricist principles almost identical to those of Locke and Condillac.[368] Thus Helvétius employs the same methodology as his predecessors, approaching the question how we may gain an accurate understanding of both morality and empirical reality by describing how human beings acquire knowledge in the first place. This is the reason why Helvétius, who appears to be primarily interested in politics and morality, nonetheless starts both his main treatises with discussions of epistemology and psychology. Helvétius thereby presents a simplified version of empiricism largely adapted from Locke and Condillac. Thus like his predecessors, Helvétius maintains that all ideas derive from sensation, stating that the mind is a *tabula rasa* at birth. Helvétius also thinks that the mind is completely devoid of innate ideas or cognitive faculties – thereby continuing the trend already visible in Condillac of eradicating all innate properties from the mind. Finally, Helvétius also follows his empiricist predecessors with his hedonistic account of human motivation and theory of the good.

Helvétius was far from the only French Enlightenment thinker to adhere to these empiricist principles. In fact, many *Philosophes* as well as the moderate Jesuits in control of the universities embraced a form empiricism inspired by Locke's *Essay*.[369] Therefore perhaps surprisingly, Helvétius's writings were almost universally criticised, rejected or even ridiculed upon publication by *Philosophes* and conservatives alike. Thus in 1758 Helvétius's first and most famous publication entitled *De l'esprit* caused a scandal that ended with the public burning and proscription of the work, while Helvétius himself only ensured his personal safety through his connections at court. Meanwhile, *De l'esprit* was refuted in the more conservative journals, whereas Helvétius's fellow

[368] David Wootton, 'Helvétius, From radical Enlightenment to Revolution', *Political Theory* 28 (2000) 307-336, 310 also emphasises the ultimately Lockean origin of Helvétius's principles.

[369] On the Jesuit acceptance of Locke's epistemology Cf. D.W. Smith, *Helvétius, A Study in Persecution* (Oxford 1965) 103-114; Jeffrey D. Burson, *The Rise and Fall of Theological Enlightenment: Jean-Martin de Prades and Ideological Polarization in Eighteenth-Century France* (Notre dame 2010) 33-54.

philosophes declined to come to his aid.[370] Partly the scandal over *De l'esprit* had been caused by the circumstances of its publication. Rather than publish the controversial work clandestinely, Helvétius had deceived an inexperienced censor into granting permission for official publication, which was promptly withdrawn when the contents of the book became more widely known. Helvétius's actions not only upset the authorities, it also angered the *Philosophes*. The scandal over *De l'esprit* called attention to the incendiary nature of philosophical publications, right at a time that the *Encyclopédie* was in danger of being censored for its controversial positions.[371] Madame du Deffand, a well-connected socialite and hostess of a salon frequented by several *Philosophes*, expressed this angle in her remark that Helvétius had upset many because he had revealed everyone's secret.[372] Deffand's remark confirms the view that Helvétius's views were widely discussed in the private environment of the salon, but that they could not be fully expressed publicly without upsetting the authorities.

Yet the uproar over Helvétius's treatises was most likely not the consequence of any theoretical innovations, but rather of the subversive purpose of his philosophy. Thus Helvétius's philosophical works are interspersed with critiques of both ecclesiastical and royal authorities, as well as contemporary moral conventions. For this reason, it would appear obvious to interpret his thought not primarily in a philosophical context, but rather in the context of contemporary debates on political and social issues. We would then conclude that Helvétius's primary intention was to criticise the politics of the *Ancien régime* and the authority of the Catholic Church. This reading would see Helvétius's philosophy – including his theory of human nature and ethics – mainly as a prelude to his long diatribes against prejudice and ignorance. His philosophy would then be portrayed not primarily as an autonomous contribution to intellectual discussion, but rather as the conceptual framework to an intervention in contemporary public debate on political and social issues. In fact, there have been several scholars who have interpreted Helvétius using such a contextual

[370] During the previous major intellectual affaire in France, caused by the doctoral thesis of the Abbé des Prades, Diderot decided to assist the author in defending his views by writing the *Suite de l'Apologie de M. l'abbé de Prades*. Yet in the case of Helvétius, Diderot decided to refrain from any action.

[371] Smith, *Helvétius* 50-51, 150.

[372] Mark Hulliung, *The Autocritique of Enlightenment: Rousseau and the Philosophes* (New Brunswick 2014 [1994]) 5; Cf. Arthur Wilson, *Diderot* (Oxford 1972) 310.

approach.[373] Yet even if our own discussion of Helvétius will not entirely ignore the social and political debates in which he sought to intervene, our primary aim will be to provide an interpretation of Helvétius philosophical ideas. We will then see that Helvétius's philosophy functions not merely as a theoretical prelude to his polemical writings, but in fact provides the basis for many of his moral and political views.

Helvétius derived this theory of human understanding and motivation almost entirely from other empiricist philosophers – in particular Locke and Condillac. Yet we will see that while Helvétius derived his fundamental principles from his empiricist predecessors, his account of human nature lacked much of the nuance and sophistication of the theories of Locke, Hobbes, and Condillac. To begin with, Helvétius simplifies Locke's axiom that all ideas are constructed from simple ideas acquired through sensation or reflection into the claim that all of our ideas are simply acquired externally – thereby seemingly denying that the mind itself has considerable liberty in how it composes it complex ideas and mixed modes. Helvétius therefore presents the mind as little more than a passive receptacle of ideas, rather than an autonomous entity capable of thinking or reasoning. In addition, Helvétius ignores the account of practical deliberation presented by Hobbes and Locke and consequently depicts human beings as determined immediately by their desire for pleasure and aversion of pain. Consequently, we will see that Helvétius's psychology and epistemology present a reductionist account of human nature, depicting man as entirely determined by his passions and environment – an account that was criticised by several of Helvétius's fellow *Philosophes*, including Diderot, Rousseau and Voltaire.

The main objective of our present discussion will then be to outline the ways in which Helvétius's moral and political philosophy follows from his theory of human nature. We will then start our investigation by examining Helvétius's theory of human nature, devoting particular attention to his conception of the mind and his account of human motivation. Throughout this discussion, we will highlight some of the similarities and differences between Helvétius's philosophy and that of his empiricist predecessors. In addition, we will present several points of critique voiced against Helvétius by his contemporaries. Subsequently, we will be able to trace the consequences of Helvétius's account of human nature

[373] Smith, *Helvétius*.

for his moral and political philosophy. In the second part of this chapter, we will thus explore how Helvétius's hedonistic conception of the Good not only leads him to a consequentialist conception of morality, but arguably to an early formulation of the utilitarian position in ethics. Finally, we will also see that following his adoption of a reductionist form of empiricism according to which the mind is nothing but a collection of ideas acquired by means of either experience or language, Helvétius concurs with Condillac and Rousseau that the transformation of morality and society requires the reform of both education and the conventional signification of moral terms.

Sensibility and Ideas

Like many of his contemporary *philosophes*, Helvétius constructs his philosophy upon three fundamental Lockean notions: the sensory origin of ideas, the mind as a *tabula rasa* and the hedonistic theory of human motivation. In fact, Helvétius sometimes refers explicitly to his authoritative predecessor to account for his positions.[374] Yet we will see that Helvétius's subsequent theories of epistemology, psychology and morality both simplified and revised Locke's empiricism. The most important of these revisions of Locke's theory had been prefigured by Condillac, whom Helvétius had undoubtedly read, but declines to mention in his writings.[375] Helvétius thus embraces Condillac's view that the mind is completely devoid of an innate organisation. Taking the notion of the mind as a *tabula rasa* as his fundamental proposition, Helvétius maintains that minds are originally completely identical, and that all intellectual differences between individuals are the consequence of the accumulation of diverging experiences.[376] Although a materialist, Helvétius denies that the physical organisation of the brain affects our cognitive faculties – unlike contemporary materialists like Diderot and d'Holbach.[377] Following Condillac, Helvétius subsequently maintains that the mind, completely devoid of any inborn characteristics, possesses three fundamental abilities. Thus the mind naturally has the ability to perceive the

[374] Helvétius, *De l'esprit* ed. by Jacques Moutaux (Paris 1988) I.iv p.47.

[375] Smith, *Helvétius* 13; 109.

[376] Helvétius, *De l'esprit* III.iv p.260.

[377] Helvétius, *De l'esprit* III.iv p.260; III.ii p.235; *De l'homme* Introduction ii t.1 p.46; II.i t.1 p.140-141. Thus d'Holbach, *Système de la nature ou les lois du Monde Physique et du Monde Moral* ed. by Jean-Pierre Jackson (Paris: Coda 2008 [1770]) I.ix p.90 and Diderot, *Refutation d'Helvétius* in: *Oeuvres completes* XXIV ed. by H. Dieckmann and J. Varloot (Paris 2004) p.540 do maintain that physical organisation affects man's cognitive abilities.

world through sensations, while its experiences are necessarily attended by feelings of pleasure and pain. Helvétius agrees with Hobbes, Locke and Condillac that the desire for pleasure and aversion of pain provides the impetus for all thought and action. Secondly, the mind can distinguish the similarities and differences between ideas, thereby giving rise to what Condillac referred to as judgment. Thirdly, the mind has the ability to memorise and recall its experiences.

Subsequently Helvétius claims that all our cognitive processes and faculties are the consequence of these three fundamental abilities of the mind: sensation, memory and the feeling of pleasure and pain. Dismissing the notion that humans possess innate cognitive faculties, Helvétius sees the mind as nothing more than the collection of its ideas derived from experience, stating: 'Qu'est-ce dans l'homme que l'esprit ? l'assemblage de ses idées.'[378] Helvétius therefore simplifies the commonplace empiricist principle that human understanding does not extend beyond things of which we may have determined ideas into the notion that the mind itself is nothing but the collection of ideas acquired through sensation.[379] At the same time, Helvétius describes all thought processes as nothing more than the perception of differences and similarities between multiple ideas retained by the mind:

> Ce principe posé, je dis encore que c'est dans la capacité que nous avons d'apercevoir les ressemblances ou les différences, les convenances ou les disconvenances qu'ont entr'eux les objets divers, que consistent toutes les opérations de l'Esprit. Or cette capacité n'est que la sensibilité physique même : tout se reduit donc à sentir.[380]

Helvétius thereby adopts the principle posed by Locke and Condillac that all mental operations ultimately consist of the perception of relations between ideas by comparing them and discovering their similarities and differences.[381] Yet in Helvétius's view, there is no essential difference between having a single sensation and perceiving two ideas at the same time while paying attention to

[378] Helvétius, *De l'homme* V.ii t.1 p.461. Cf. II.ii t.1 p.149 ;

[379] *De l'esprit* I.i p.21 : 'Pour nous assurer de cette vérité, considerons la nature. Elle nous présente des objets ; ces objets ont des rapports avec nous et des rapports entr'eux ; la connoissance de ces rapports forme ce qu'on appelle l'*Esprit* : il est plus ou moins grand, selon que nos connoissances en ce genre sont plus ou moins étendues.'

[380] Helvétius, *De l'esprit* I.i p.21.

[381] Helvétius, *De l'esprit* I.i p.22; *De l'homme* II.iv t.1 p.157-158. Cf. D'Alembert, *Essai sur les éléments de philosophie* ed. by Richard Schwab (Hildesheim 1965) V p.74, 80-82.

their differences and similarities.[382] Consequently, Helvétius maintains that thinking or judging is ultimately not different from sensation itself, concluding that 'toutes les opérations de l'esprit se réduisent à des pures sensations.'[383]

Yet for Helvétius, human sensibility not only enables judgment of the similarities and differences between ideas, but also includes the experience of feelings of pleasure and pain necessarily conjoined with sensations. Like Condillac, Helvétius maintains that these feelings rouse the otherwise mind passive mind into action, and have a defining influence on how it conducts its operations:

> Il résulte de ce Chapitre que tous les jugemens occasionnés par la comparaison des objets entr'eux, supposent en nous intérêt de les comparer. Or cet intérêt nécessairement fondé sur l'amour de notre bonheur, ne peut être qu'un effet de la sensibilité physique, puisque toutes nos peines et nos plaisirs y prennent leur source. Cette question examinée, j'en conclurai que la douleur et le plaisir physique est le principe ignoré de toutes les actions des hommes.[384]

Following Condillac, Helvétius claims that without a personal interest derived from feelings of pleasure and pain, the mind is at a complete standstill.[385] Personal interest therefore directs our attention, as the mind naturally focuses on ideas most relevant to its purpose to attain pleasure and avoid pain. But our desire for pleasure and aversion of pain also produces the will. Clearly following Hobbes, Helvétius then states that it does not make sense to apply the word 'liberté' to the notion of 'volonté,' as the will is always causally determined by the interaction between passions and sensations.[386] Yet Helvétius does not reiterate

---

[382] Helvétius, *De l'esprit* I.i p.22-25; *De l'homme* II.iv t.1 p.157-159: 'Cela pose, toutes les opérations de l'esprit se réduisent à des pures sensations. Pourquoi donc admettre en nous une faculté de juger distincte de la faculté de sentir ?'

[383] Helvétius, *De l'homme* II.iv t.1 p.157-159. Cf. Helvétius, *De l'esprit* I.i p.24. As we have seen, Hobbes makes a similar assertion in *De corpore* VI.xxv.8 p.399.

[384] Helvétius, *De l'homme* II.vi t.1 p.170. Cf. III.iv t.1 p.319.

[385] Helvétius, *De l'homme* II.vi t.1 p.167. Cf. D'Holbach, *Système de la nature* I.vi p.52-53.

[386] Helvétius, *De l'esprit* I.iv p.47: 'On a donc une idée nette de ce mot de *Liberté*, pris dans une signification commune. Il n'en est pas ainsi lorsqu'on applique ce mot de *liberté* à la volonté. Que seroit ce alors que la liberté ? On ne pourroit entendre, par ce mot, que le pouvoir libre de vouloir ou de ne pas vouloir une chose ; mais ce pouvoir supposeroit qu'il peut y avoir des volontés sans motifs, et par conséquent des effets sans cause. Il faudroit donc que nous pussions également nous vouloir du bien ou du mal ; supposition absolument impossible. En effet, si le desir du plaisir est le principe de toutes nos pensées et de toutes nos actions, si tous les hommes tendent continuellement

Hobbes's extensive account of practical deliberation that describes how human beings are motivated by desires and aversions, but may nonetheless choose to ignore certain passions on account of their foreseen consequences. Helvétius in fact claims that personal interest has such a strong impact on mental operations that it directly impacts our judgment.[387] Thus Helvétius thinks that as the mind does not have any autonomy in judgment, its perception of the relations between ideas may often be determined by its passions and interests. Self-interest even makes men contradict self-evident truths, such as the maxim 'do unto others what you would have them do to you.'[388] This claim therefore leads Helvétius to the conclusion that passions not only instigate thinking, they also influence the outcome of our judgments. As Helvétius thus portrays human beings as substantially and inevitably directed by their passions in both thought and action, he relinquishes the notion that humans may transcend their immediate passions and identify with the common good through their capacity for reasoning.[389]

The Mind as Passive

With this conception of human understanding as a collection of ideas endowed with sensibility, Helvétius discards the last vestige of the notion that the mind is an active and autonomous entity.[390] Thus we have seen that while Locke maintained that the mind solely acquires its simple ideas through sensation or reflection, he nonetheless assumed that the mind is active, and has considerable liberty in the way it structures simple ideas into more complex aggregates and then connects these to each other. This assumption also applies to mixed modes – the type of ideas used in moral reasoning. Locke therefore considered the mind an entity independent from both its ideas and sensations. Yet despite Locke's extensive account of the operations of the understanding, he never truly explains the origin or properties of this autonomous mind. Consequently, Locke seems to

---

vers leur bonheur réel ou apparent, toutes nos volontés ne sont donc que l'effet de cette tendance.' Cf. D'Holbach, *Système de la nature* I.vi p.53; Hobbes, *Leviathan* I.v.5; I.vi.53-54; II.xxi.1-4.

[387] Helvétius, *De l'homme* IX.xv t.2 p.809.

[388] Helvétius, *De l'homme* IX.xxiii t.2 p.829.

[389] Michael Sonenscher, *Before the Deluge: Public Debt, Inequality, and the Intellectual Origins of the French Revolution* (Princeton 2007) 275 notes that Helvétius' reductionist conception of the mind as determined entirely by sensation and self-interest, in fact leads him to emphasise the importance of public debate and freedom of expression. As he denies that humans are able to transcend their own perspective and self-interest by means of reasoning, Helvétius thinks that the only way to identify the common good is through debate in the public sphere.

[390] Helvétius, *De l'esprit* I.i p.24. Cf. D'Holbach, *Système de la nature* I.vi p.55.

retain an almost Cartesian notion of the mind as a thinking thing, but without openly acknowledging or explaining its role in higher cognitive functions.[391] In the previous chapter, we have seen that Condillac had already dismissed Locke's implicit assumption that the mind is an autonomous entity with innate cognitive faculties as incoherent. Instead, Condillac maintained that our cognitive operations are nothing but habits developed through sensation and the use of language.

With his conception of the mind as nothing but a collection of ideas endowed with sensibility, Helvétius thus follows Condillac in the rejection of Locke's tacit assumption that the mind would have innate cognitive faculties. Yet Helvétius does not adopt Condillac's alternative explanation of man's cognitive faculties as developed through language. To be sure, we have seen in the previous chapter that Condillac's theory is not without its own problems. But at least Condillac's account of the cognitive uses of language provides some explanation of how human understanding can advance beyond the level of mere instinct. By contrast, Helvétius's theory of language is limited to the recognition that moral concepts are tied to terms with a conventional signification. Accordingly, Helvétius presents a very reductionist theory of the human mind as nothing more than a collection of ideas endowed with sensibility – a theory sometimes classified by modern scholars as a form of 'sensationism.'[392] Yet without Condillac's account of the role of language in cognition, Helvétius's theory of the mind has difficulty explaining how human beings can possess the higher cognitive faculties that distinguish them from animals.

---

[391] Diderot, *D'Alembert's Dream* trans. by Leonard Tancock (London: Penguin 1966)152-162 aptly illustrates the perplexity of this position by comparing the mind to a piano that plays itself or a book that can read its own contents. Yet Diderot does not necessarily claim thereby that for this reason this conception of the mind is untenable – merely that it does not resemble any other property of experience.

[392] O'Neal, *The Authority of Experience: Sensationalist Theory in the French Enlightenment* (Penn University Park 1996). The terms 'sensationism' or 'sensationalism' are often used by scholars to denote a type of empiricist philosophy that intends to reduce all mental operations to man's capacity for sensibility, thereby rejecting the notion of innate cognitive faculties. In this sense, Helvétius' philosophy would clearly qualify as a species of sensationism. Yet as is often the case with these kinds of epithets, the term is generally used imprecisely to describe various kinds of theories. Thus Condillac and even Locke are sometimes classified as sensationists. Because in this thesis we aim to highlight the differences between various empiricist thinkers, we will refrain from using this kind of generalising terminology.

Helvétius nonetheless attempts to render his theory more convincing by providing alternative explanations for certain properties of human nature. Firstly, in view of his claim that the mind is not free to make its own judgments, but instead merely senses the similarities and differences between ideas, Helvétius must explain how it is possible that humans unquestionably and continuously err in their judgments. Helvétius considers errors of judgment not as mistakes in reasoning made by an autonomously reflective mind, but rather as the accidental cause of external factors:

> J'ai montré les vraies causes de nos faux jugemens; j'ai fait voir que toutes les erreurs de l'esprit ont leur source ou dans les passions, ou dans l'ignorance, soit de certains faits, soit de la vraie signification de certains mots. L'erreur n'est donc pas essentiellement attachée à la nature de l'esprit humain ; nos faux jugemens sont donc l'effet de causes accidentelles, qui ne supposent point en nous une faculté de juger distincte de la faculté de sentir.[393]

Claiming that errors do not originate in the nature of the mind itself, Helvétius identifies three external causes for its mistakes. Thus faulty conclusions may simply be caused by an insufficient grasp of the facts of empirical reality, and therefore by ignorance rather than error of judgment. In addition, Helvétius claims that the passions may subvert the accurate perception of the relations between ideas – a position we have already encountered in our discussion of the role of self-interest in cognition. Thirdly, following Condillac's exploration of the role of language in cognition, Helvétius agrees that errors of judgment may also be caused by ignorance of the proper signification of words. As we will see in our subsequent discussion of Helvétius's ethical theory, the role of language as a source of error is especially significant to reflections on morality – a subject entirely conducted through concepts tied to linguistics terms without a direct connection to empirical reality.

Secondly, Helvétius must concede that, despite his conception of the mind as devoid of any innate characteristics, individuals nonetheless display a distinct disparity in their intellectual abilities. As he denies that these disparities may be caused by differences in physical or mental organisation, Helvétius maintains that this disparity can only be explained by the fact that all minds are

[393] Helvétius, *De l'esprit* I.iv p.50.

comprised of different ideas, as no two individuals experience entirely the same sensations:

> D'où provient donc l'extrême inégalité des esprits ? de ce que personne ne voit précisément les mêmes objets ; ne s'est précisément trouvé dans les mêmes positions ; n'a reçu la même éducations ; et de ce qu'enfin le hazard qui préside à notre instruction ne conduit pas tous les hommes à des mines également riches et fécondes. C'est donc à l'éducation prise dans toute l'étendue du sens qu'on peut attacher à ce mot, et dans lequel même l'idée du hazard se trouve comprise qu'on peut rapporter l'inégalité des esprits.[394]

Rejecting any form of innatism, Helvétius claims that all the manifest differences between individuals – in intellectual capacities as well as personal character – are solely the consequence of differences in the collections of ideas that comprise their minds. As Helvétius thinks that the mind itself consists of nothing but the ideas it acquires through experience, he concludes that there is only one method to compose a sound mind: acquire the right ideas. Subsequently, we will see that this conclusion is at the basis of Helvétius's insistence on the importance of education.

Ultimately, Helvétius simplifies the main tenets of previous empiricist thinkers to the extent that his philosophy arguably no longer presents a satisfactory theory of human nature.[395] Following Helvétius's repudiation of any form of cognitive autonomy, it is questionable whether according to his theory human beings could still be said to think or reason. In addition, Helvétius accords the passions an inescapable influence over human thought and action. Consequently, Helvétius depicts the mind as nothing more than a collection of ideas, while its cognitive processes and abilities are nothing more than the blind interaction of its ideas, set in motion and even controlled by its passions and sentiments. According to this conception of the mind, human nature is thus entirely determined by instinct. In many ways, Helvétius's conception of man is reminiscent of the sensitive statue depicted in Condillac's *Treatise on Sensations*, or of human nature as described in de la Mettrie's *L'homme machine*. In fact, Helvétius himself invites this comparison when he states that: 'L'Homme est une

[394] Helvétius, *De l'homme* IV.xxiv t.1 p.418.

[395] This assessment is shared by Sophie Audidière, 'Why do Helvétius's writings matter? Rousseau's *Notes sur De l'esprit*', *British Journal for the History of Philosophy* 24 (2016) 983-1001, 998.

machine qui mise en mouvement par la sensibilité physique doit faire tout ce qu'elle exécute.'[396]

Self-Interest and Morality

In the preceding section, we have seen that Helvétius's epistemology discards all autonomous faculties of the mind and claims that all cognitive processes are occasioned by the interaction of ideas acquired through experience and directed by the passions. Accordingly, Helvétius follows Locke in rejecting the notion that human beings would have some innate faculty of reason that would allow them to distinguish right from wrong. Yet Helvétius also explicitly dismisses the idea – first mentioned by Shaftesbury and subsequently developed by Hutcheson, Rousseau and Hume among others – that human sensibility not only includes sensations of empirical objects imbued with pleasure and pain, but also certain moral sentiments.[397] Instead, Helvétius constructs his entire system of ethics on the only remaining normative element of human nature – the interest to attain pleasure and avoid pain:

> Si l'Univers physique est soumis aux lois du mouvement, l'Univers moral ne l'est pas moins à celles de l'intérêt. L'intérêt est sur la terre le puissant enchanteur qui change aux yeux de toutes les Créatures la forme de tous les objets.[398]

According to Helvétius, humans are ingrained with such an unavoidable tendency to promote their self-interest that any theory of morality seeking to regulate their actions would have to take this tendency as its starting point. Helvétius subsequently attempts to demonstrate that all our actions can indeed be explained as the consequence of the interest to attain pleasure and avoid pain.[399] For instance, in order to explain why humans sometimes display selfless behaviour ostensibly motivated by compassion or benevolence, Helvétius claims that even these sentiments are ultimately nothing but a modification of our own self-interest, as we expect our benevolent actions to be rewarded in some way or another. Helvétius thus takes a rather cynical view of human nature, as even the

[396] Helvétius, *De l'homme* II.x p. t.1 194.
[397] Helvétius, *De l'homme* V.iii t.1 p.465. Cf. D'Holbach, *Système de la nature* I.x p.108.
[398] Helvétius, *De l'esprit* II.ii p.59.
[399] Helvétius, *De l'homme* IV.xxii t.1 p.411.

most noble and altruistic acts, including parental love, turn out to be attempts to satisfy our passions and desires.[400]

The hedonistic account of human motivation is a preconception common to many enlightenment thinkers. Thus some scholars have described Helvétius's hedonism as part of a much broader revival of the Epicurean position in ethics, of which La Rochefoucauld, Gassendi, Mandeville, and La Mettrie are often considered the primary proponents.[401] Yet some of Helvétius's contemporaries considered his uncompromising position that *all* human thoughts and actions are occasioned by a desire for pleasure or aversion of pain as too reductionist an account of human motivation. Thus Diderot, who agreed that human beings are often motivated by their desires and aversions, nonetheless called out Helvétius on his simplistic theory of human motivation:

> Hé bien, Mr. Helvétius, tous les projets d'un grand roi, toutes les fatigues d'un grand ministre ou d'un grand magistrat, toutes les méditations d'un politique, d'un homme de génie, se réduisent donc à foutre un coup le matin et à faire un étron le soir. Et vous appelez cela faire de la morale et connaître l'homme.[402]

Diderot's objection is that despite the fact that individuals are commonly motivated by desires or aversions, not all human behaviour can be explained as directly their consequence. For Diderot, some thoughts or actions are not simply the product of immediate self-interest, but are occasioned by ulterior motivations, such as the love of beauty, truth, virtue or justice. Diderot further explores this question of human motivation in several of his other works, such as in the *Le neveu de Rameau*, a fictional dialogue between a high-minded philosopher with a love of truth and virtue, and the nephew of the famous composer – a scoundrel who sees the pursuit of pleasure as his sole objective.[403]

---

[400] Helvétius, *De l'esprit* IV.x p.491-493.

[401] Pierre Force, 'Helvétius as an Epicurean political theorist', in: Neven Leddy and Avi S. Lifschitz (eds.), *Epicurus in the Enlightenment* (Oxford 2009). All these authors accepted the principle that human beings are governed by their passions and desires, even though they did not all accord self-interest the same function in their psychology and moral philosophy. Yet in previous chapters we have seen that many more authors, including Hobbes, Locke, and Condillac, equally embraced a hedonistic account of human motivation – raising the question whether by the early modern era the position could still be described as a form Epicureanism.

[402] Diderot, *Refutation d'Helvétius* p.529. Cf. Jorn Schosler, 'Rousseau et Diderot, critiques de la philosophie égalitaire d'Helvéius', *Revue romane* 15 (1980) 68-83 for a more extensive discussion of Diderot's critique of Helvétius.

[403] Diderot, *Rameau's Nephew* trans. by Leonard Tancock (London 1966) 64-65.

Yet as neither interlocutor succeeds in presenting a convincing case for their view of human nature, the dialogue ends in aporia – perhaps signifying that Diderot himself would not take a definitive position on the question of the origins of human motivation.[404]

As Hobbes, Locke, and Condillac also ascribed to a hedonistic account of human motivation, Diderot's objection may also apply to their moral philosophy. Indeed, David Hume raised an objection similar to that of Diderot against what he perceived as Hobbes's account of human motivation.[405] Yet beyond their similarity in presuppositions, there is an important difference between Helvétius's theory and the moral philosophy of either Hobbes or Locke. Thus in previous chapters, we have seen that both Hobbes and Locke ascribed to individuals the ability for careful deliberation, by which a person may ascertain whether it would be prudent to pursue or ignore a certain passion. Thus according to Hobbes, human beings may use their ability for reasoning – which is itself developed through the use of language – to ascertain the long-term consequences of their actions. Although we have seen that Locke's account of moral reasoning is less straightforward, he too ascribes to human beings the ability to determine the prudence of their actions. This capacity for prudential reasoning not only ascribes to human beings the ability to go beyond personal experiences and immediate desires by reflecting on their long term interests, but also to identify a rule of morality that would safeguard the well-being of themselves as well as society as a whole.

Yet in view of Helvétius's theory of the understanding, it is not clear to what extent he would ascribe to human beings the same ability for practical

---

[404] Macintyre, *After Virtue* 56 presents Diderot's dialogue as an example of why the modern development towards consequentialist or utilitarian conceptions of ethics falls short, writing that: 'The challenge that Rameau presents to the *philosophe* cannot of course be met within the terms of Diderot's own thought. For what divides them is the question of precisely which of our desires are to be acknowledged as legitimate guides to action, and which on the other hand are to be inhibited, frustrated or re-educated; and clearly this question cannot be answered by trying to use our desires themselves as some sort of criterion.'

[405] Hume, *Enquiry concerning the Principles of Morals* appendix II §4 p.91: 'An EPICUREAN or a HOBBIST readily allows, that there is such a things as friendship in the world, without hypocrisy or disguise; though he may attempt, by a philosophical chymistry, to resolve the elements of this passion, if I may so speak, into those of another, and explain every affection to be self-love, twisted and moulded, by a particular turn of imagination, into a variety of appearances.' [...] 'All attempts of this kind have hitherto proved fruitless, and seem to have proceeded entirely, from that love of *simplicity*, which has been the source of much false reasoning in philosophy.'

deliberation and moral reasoning. This issue first emerges in his account of the origin of society. Evoking Hobbes's account of the state of nature, Helvétius maintains that before the establishment of society, all individuals seek to satisfy their passions and desires without regard for each other, leading to widespread conflict.[406] In the absence of laws, there is no rule of justice and consequently 'Ce n'est ni l'équité, ni même l'apparence de l'équité qui juge entre le foible et le puissant ; mais la force, le crime et la tyrannie.'[407] As Helvétius claims that in the state of nature the right of the strongest reigns, he criticises the claim that human beings have a natural tendency for sociability as unintelligible and speculative.[408] But Helvétius then maintains that sociability will nonetheless develop out of the concern for self-interest of individuals:

> L'intérêt et le besoin sont le principe de toute sociabilité. Ce principe (dont peu d'Ecrivains ont donné des idées nettes) est donc le seul qui unisse les hommes entr'eux. Aussi la force de leur union est-elle toujours proportionnée à celle et de l'habitude et du besoin.[409]

Echoing Pufendorf's as well as Condillac's accounts of the origin of sociability, Helvétius maintains that at some point humans in the state of nature realised that their self-interest would be better served if they pursued it in coordination rather than in competition with each other.[410] It is this insight that is at the basis of the emergence of both morality and the state.

Yet upon Helvétius's own theory of human understanding, it is not clear how this insight would occur to man in the state of nature, without any previous experience of society. First of all, Helvétius's theory lacks an account of practical deliberation by which human beings can foresee the long-term consequences of their actions. As we have seen, Helvétius's conception of human nature is similar

---

[406] Helvétius, *De l'esprit* III.iv p.250.

[407] Helvétius, *De l'homme* IV.x t.1 p.360. Cf. IV.viii t.1 p.351-353;

[408] Helvétius, *De l'homme* II.viii t.1 p.183-184.

[409] Helvétius, *De l'homme* II.viii t.1 p.182.

[410] Pufendorf, *On the Duty of Man and Citizen* I.iii.7 p.35. Pufendorf's account of sociability had become commonplace during the eighteenth century. E.g. Cf. Voltaire, *Letters on England* ed. and trans. by Leonard Tancock (London 1980) 128: 'It is as impossible for a society to be formed and be durable without self-interest as it would be to produce children without carnal desire or to think of eating without appetite, etc. It is love of self that encourages love of others, it is through our mutual needs that we are useful to the human race. That is the foundation of all commerce, the eternal link between men.'

to that of Condillac's animated statue – a being immediately impelled by its desires and aversions. In the absence of the ability for long-term prudential reasoning, it is unclear how such a being could foresee the benefits of society without ever having experienced sociable existence. Secondly, Helvétius's theory of the understanding also fails to explain how human beings are capable of devising moral concepts – ideas not directly acquired through experience. Consequently, it is difficult to see how Helvétius could explain that human beings in the state of nature acquire or construct the various moral concepts required for the institution of society. In view of these deficiencies, Helvétius's account of the origin of society is especially susceptible to Rousseau's critique, discussed more extensively in our next chapter, that moral philosophers often presuppose civilised man, in possession of extensive foresight as well as moral concepts, for natural man in their account of the state of nature.

Self-Interest and Utilitarianism

After considering the origin of society, Helvétius turns to the question of the purpose of the state. Concerning this issue, Helvétius again follows Hobbes at first, but nonetheless introduces a novel element into his political theory. Thus starting out from the assumption that all humans seek to satisfy their desires, Hobbes maintains that the primary purpose of the sovereign is to regulate conflicts among citizens that result from the self-interested pursuit of their desires. Helvétius does not dispute Hobbes's point, stating that public interest requires citizens 's'enchaîner par des Loix sages, et se mettre dans l'heureuse impuissance de se nuire.'[411] This legislation would then not merely dole out punishments for the self-serving, it would also reward self-sacrifice in service of the common good.[412] Yet besides restraining and rewarding citizens, Helvétius thinks that laws should also provide society with favourable circumstances for its inhabitants:

> Des Loix sages pourroient sans doute opérer le prodige d'une félicité universelle. Tous les Citoyens ont-ils quelque propriété ? Tous sont-ils dans un certain état d'aisance,

[411] Helvétius, *De l'homme* IV.xxi t.1 p.407.

[412] Helvétius, *De l'esprit* III.iv p.251: 'La peine et la récompense sont les deux seuls liens par lesquels ils ont pu tenir l'intérêt particulier uni à l'intérêt général.'

> peuvent-ils par un travail de sept ou huit heures subvenir abondamment à leur besoins et à ceux de leur famille ? Ils sont aussi heureux qu'ils peuvent l'être.[413]

For Helvétius, the role of morality and the state is thus not merely to provide security by preventing and mediating conflicts among citizens. The state is also expected to take a more active role in securing the happiness of its citizens – for instance by securing an even distribution of wealth. Thus Helvétius states that the immediate objective of morality and the state is to secure 'le bonheur du plus grand nombre.'[414]

By adopting this principle, Helvétius clearly prefigures Jeremy Bentham's famous principle that the objective of morality is 'the greatest happiness of the greatest number.'[415] Some historians of moral philosophy have therefore identified Helvétius as Bentham's precursor and thereby as one of the first thinkers to formulate a consistent and unmitigated theory of utilitarianism – even if these historians often do not discuss his philosophy at length and focus on the development of the theory in Britain during the late eighteenth and nineteenth centuries.[416] In fact, the similarities between the two eighteenth-century utilitarian thinkers had already been recognised by John Stuart Mill, who noted that 'the premises of Bentham are all clearly given by Helvétius.'[417] There is certainly a kernel of truth in this analysis – albeit with two caveats. First of all, the difference between Helvétius's utilitarianism and Hobbes's position on the role of the state is in emphasis rather than principle. Thus Hobbes would not have denied that ultimately, the function of the sovereign is to secure the well-

[413] Helvétius, *De l'homme* VIII.i t.2 p.659.

[414] Helvétius, *De l'homme* IX.vi t.2 p.773.

[415] As Stephen Darwall, 'The Foundations of Morality', 234 and 239 points out, this principle was not altogether new, as it had already been formulated by Leibniz and Hutcheson, albeit as part of fundamentally different ethical systems. Joachim Hruscka, 'The Greatest Happiness Principle and Other Early German Anticipations of Utilitarian Theory', *Utilitas* 3 (1991) 165-177 identifies a variety of German natural lawyers from the late seventeenth and early eighteenth centuries who likewise promoted the principle.

[416] Heydt, 'Utilitarianism before Bentham', 30-33; Rosen, *Classical Utilitarianism from Hume to Mill* 83 and 95; Force, *Self-Interest before Adam Smith* 93-94; Hirschmann, *The Passions and the Interests* 27-28; Everett C. Ladd, Jr., 'Helvétius and D'Holbach: "La Moralisation de la Politique"', *Journal of the History of Ideas* 23 (1962) 221-238, 224 even calls Helvétius 'the founder of doctrinal utilitarianism.'

[417] Quoted in Force, 'Helvétius as an Epicurean political theorist', 118. Original passage in: John Stuart Mill, 'Bentham' *London and Westminster Review* August 1838.

being of its subjects. Helvétius merely accords the state a more active role in guaranteeing the happiness of its citizens.

Secondly, Helvétius was not the only eighteenth century French thinker to prefigure later formulations of utilitarianism. Thus many *Philosophes*, including Diderot, d'Holbach, Voltaire, Condillac and arguably even Rousseau, employed utilitarian arguments in their critique of contemporary political and social practices.[418] For example, in his *Supplément au voyage de Bougainville*, Diderot claimed that French sexual mores and marriage laws promoted neither private nor public happiness, and should therefore be revised.[419] The same method is also employed by Helvétius, who supports his endless diatribes against the *Ancien régime* and the Catholic Church with the argument that they both obstruct individuals in the pursuit of their self-interest and harm the public good. Seen in this light, Helvétius's utilitarianism could be considered as much as a convenient starting point for his critique of public affairs, as an innovation in moral philosophy. Nonetheless, there is at least one major difference between Helvétius's theory and that of contemporary *Philosophes*. Whereas the latter often employed utilitarian arguments, Helvétius was the only French thinker to take 'le bonheur du plus grand nombre' as the fundamental principle of his moral philosophy. Arguably, this is sufficient reason to designate Helvétius the first thinker to present an unambiguous formulation of moral utilitarianism.

The Reform of Society

Starting out from this utilitarian principle, Helvétius subsequently maintains that current legislation and moral convention often conflict both with the private interests of individuals and with the public good of society as a whole, lamenting that 'la morale actuelle, comme je viens de le dire, n'est qu'un tissu d'erreurs et de contradictions grossieres.'[420] The main problem is that the moral conventions currently prevalent in society are antiquated and often privilege certain groups over society as a whole. Both the law and public morality are thus full of these

[418] Pierre Rosanvallon, 'L'utilitarisme français et les ambiguités de la culture politique prérévolutionnaire (position d'un problème)', in: Keith Michael Baker (ed.), *The Political Culture of the Old Regime* I (Oxford 1987) 435-440, 437 notes that Helvétius' 'Œuvre est tout simplement le symptôme et l'expression d'un mouvement intellectuel diffus qui sous-tend toute la culture politique française de la fin du XVIIIe siècle.' Cf. Hulliung, *The Autocritique of Enlightenment* 11-44.

[419] Diderot, *Supplément au voyage de Bougainville* in: John Mason and Robert Wokler (eds. And trans.), *Political Writings* (Cambridge 1992) 52; 67.

[420] Helvétius, *De l'homme* X.vii t.2 p.903.

préjugés – moral or legal conventions that contradict Helvétius's utilitarian principle of morality as the greatest good of the greatest number. Yet unlike Bernard Mandeville, who also believed that prevailing moral conventions often pervert the common interest, Helvétius does not conclude that the notion of public morality may never be anything other than a scam, invented and promoted by skilful politicians 'that they might reap the Fruits of the Labour and Self-denial of others'[421] Mandeville maintained that in the absence of a morality of the common good, the private vices of individuals would in fact lead to public benefits, claiming that man's 'vilest and most hateful Qualities are the most necessary Accomplishments to fit him for the largest, and according to the World, the happiest and most flourishing Societies.'[422]

Helvétius agreed with Mandeville to the extent that he recognises that certain passions conventionally considered as vices in private individuals, such as the desire for luxury, may turn out to be beneficial to the public interest. Yet for Helvétius the unbridled pursuit of individual passions would certainly lead to excessive conflict among individuals and therefore subvert the public good. According to Helvétius, the goal of securing the greatest happiness for the greatest number is then accomplished not simply by restraining the passions of individuals, but rather by redirecting them towards the public good:

> Je dis que tous les hommes ne tendent qu'à leur bonheur ; qu'on ne peut les soustraire à cette tendance ; qu'il seroit inutile de l'entreprendre, et dangereux d'y réussir ; que par conséquent, l'on ne peut les rendre vertueux qu'en unissant l'intérêt personnel à l'intérêt général.[423]

The purpose of morality and legislation is thus to harness the passions of individuals and align their private interests with the common good, leading to self-interested behaviour that nonetheless benefits society as a whole.[424] Far from denying the necessity of virtue and morality, Helvétius thus maintains that established morality and legislation should be reformed, cleansing it of

[421] Mandeville, *An Enquiry into the Origin of Moral Virtue* in: idem, *The Fable of the Bees* ed. by Philip Harth (London 1989) 86. On Mandeville and his influence on Enlightenment moral philosophy cf. E.J. Hundert, The Enlightenment's Fable: Bernard Mandeville and the Discovery of Society (Cambridge 2005).

[422] Mandeville, *The Preface* in: idem, *The Fable of the Bees* ed. by Phillip Harth (London 1970) 53.

[423] Helvétius, *De l'esprit* II.xv p.152.

[424] Helvétius, *De l'esprit* II.xiii p.133; II.vi p.81-84. Cf. D'Holbach, *Système de la nature* I.ix p.89.

accumulated prejudice favouring the interests of factions and re-aligning it with the public good.[425] Consequently, Helvétius states that: 'par ce mot de vertu, l'on ne peut entendre que le desir du bonheur général ; que, par conséquent, le bien public est l'objet de la vertu.'[426] Helvétius subsequently presents public debate as the most suitable method for the identification of the public interest.[427]

Yet Helvétius also thinks that a reformed utilitarian morality should be actively promoted and fostered among citizens in society. Helvétius therefore presents two proposals for the implementation of reformed morality that clearly conform to his theory of the mind as a passive entity. The first element of this plan is to educate – or indoctrinate – individuals with proper moral ideas. We have already seen that following his eradication of all innate properties from his conception of the mind, Helvétius maintains that the thoughts of human beings are merely the consequence of the perception of relations between ideas already present within the understanding. This principle equally applies to ideas of morality, which in Helvétius's view are ultimately merely conventional. The moral outlook of individuals is therefore wholly determined by ideas acquired from their social environment.[428] In Helvétius's view, the moral outlook of his contemporaries is therefore shaped by the prejudices pervading society. To remedy this situation, Helvétius maintains first of all that proper moral education is crucial for the implementation of a morality in line with the common interest. For instance, Helvétius proposes a secular catechism to be taught to school children, expounding and inculcating a morality conducive to the common good – thereby prefiguring Mably's *Catéchisme du citoyen,* published two years before the French Revolution.[429]

Yet this reform of morality through education would have to go hand in hand with a revision of moral terminology. Following Locke and Condillac, Helvétius maintains that moral reflection and judgment is conducted through the

[425] Helvétius, *De l'esprit* II.xiv p.135-146.

[426] Helvétius, *De l'esprit* II.xiii p.128. Cf. D'Holbach, *Système social* VI p.61.

[427] Helvétius, *De l'homme* IX.xii t.2 p.797: 'Veut-on en ce genre s'assurer de la vérité de ses opinions ? il faut les promulguer. C'est à la pierre de touche de la contradiction qu'il faut les éprouver. La presse doit donc être libre. La magistrat qui la gêne s'oppose donc à la perfection de la Morale et de la Politique : il peche contre sa Nation ; il étouffe jusque dans leurs germes les idées heureuses qu'eût produit cette liberté.' Cf. Helvétius, *De l'homme* IX.vii t.2 p.777

[428] This was a common view among the *philosophes*. Cf. Voltaire, *Poème sur la loi naturelle* ed. by Francis J. Crowley (Berkeley 1938) III p.268 : 'Que conclure à la fin de tous mes longs propos ? / C'est que les préjugés sont la raison des sots ; / Il ne faut pas pour eux se déclarer la guerre.'

[429] Helvétius, *De l'homme* X.vii t.2 p.903-910 ; *De l'esprit* II.xvii p.158

consideration of moral concepts attached to certain conventional terms. Furthermore, these moral ideas are peculiar in that they are not acquired directly through experience of empirical reality, but rather transmitted through language. According to Helvétius, this has caused widespread confusion in the moral concepts prevalent in society:

> Faute d'une éducation saine, les hommes n'ont de la bonté morale que des idées obscures. Ce mot *bonté* arbitrairement employé par eux, ne rappelle à leur souvenir que les diverses applications qu'ils en ont entendu faire ; applications toujours différentes et contradictoires, selon la diversité, et des intérêts et des positions de ceux avec lesquels ils vivent. Pour convenir universellement de la signification du mot *bon* appliqué à la moral, il faudroit qu'un excellent Dictionnaire en eût déterminé le sens précis.[430]

In this passage, Helvétius adopts the perspective on the role of language in cognition that was developed by Condillac out of Locke's reflections on language in book III of the *Essay*. In the preceding chapter, we have seen that Condillac worried that the imprecision and inaccuracy of linguistic conventions would affect the composition of ideas, thereby determining our reasoning. Helvétius agrees with Condillac that over the course of history, language has gradually become a defective method of communication.[431] Particularly terms regarding politics, metaphysics and morality have become flawed and imprecise in signification. Helvétius subsequently echoes Condillac's suggestion that language should be reformed wholesale, with a more modest proposal for the definition of moral terminology.

In previous chapters, we have seen that Locke and Condillac had also advised the reform of both education and linguistic conventions – in particular with regard to moral terminology. Ultimately, Helvétius's proposals for the reform of education and language do not seem fundamentally different from

[430] Helvétius, *De l'homme* II.xvi t.1 p.229. Cf. II.xvi t.1 p.227.

[431] Helvétius, *De l'esprit* I.iv p.49-50: Ce n'est point aux Philosophes, c'est au besoin qu'on doit l'invention des Langues ; et le besoin, en ce genre, n'est pas difficile à satisfaire. En conséquence, on a d'abord attaché quelques fausses idées à certains mots ; ensuite on a combiné, comparé ces idées et ces mots entr'eux ; chaque nouvelle combinaison a produit une nouvelle erreur ; ces erreurs se sont multipliées, et en se multipliant, se sont tellement compliquées qu'il seroit maintenant impossible, sans une peine et un travail infini, d'en suivre et d'en découvrir la source. Il en est des Langues comme d'un calcul algébrique : il s'y glisse d'abord quelques erreurs ; ces erreurs ne sont pas aperçûes ; on calcule d'après ces premiers calculs ; de proposition en proposition, l'on arrive à des conséquences entiérement ridicules.

those of his predecessors. Yet when placed in the broader context of his philosophy, Helvétius's proposals turn out to have greater implications. Thus we have seen that Helvétius conceives the mind as nothing more than a receptacle of ideas without any cognitive autonomy. According to Helvétius's theory, this entails that our moral outlook is entirely determined by the moral conventions prevalent in our social environment.[432] Consequently, Helvétius seems to think that humans are not naturally moral beings, but can only be made virtuous by inculcating them with moral concepts in line with the common good.[433] Unlike Hobbes and Locke, Helvétius does not claim that individuals may use their reason to perceive the moral guidelines that would lead them to felicity. Rather, Helvétius sees morality as a conventional rule disseminated through education and language that would direct citizens towards the common good. The result is a conception of society in which the state is responsible both for the elimination of prejudice and the dissemination of a morality that directs its citizens to the greatest happiness for the greatest number.

Conclusion

We have seen that Helvétius proceeds from the same presuppositions as his empiricist predecessors including Hobbes, Locke and Condillac. Consequently, Helvétius's thought displays several characteristics that we have also encountered in our previous chapters. Thus like other empiricists, Helvétius's

[432] Helvétius, *De l'esprit* I.iv p.47: 'Si, comme Mr. Locke l'a prouvé, nous sommes disciples des amis, des parens, des lectures, et enfin de tous les objets qui nous environnent ; il faut que toutes nos pensées et nos volontés soient des effets immédiats, ou des suites nécessaires des impressions que nous avons reçûës.

[433] The limitations of this conclusion was not lost on Helvétius contemporaries – among them Voltaire, who responded to the position taken by Helvétius' in his *Poème sur la loi naturelle* II p.254-255: 'On insiste, on me dit : « L'enfant dans son berceau / N'est point illuminé par ce divin flambeau ; / C'est l'éducation qui forme ses pensées ; / Par l'exemple d'autrui ses mœurs lui sont tracées ; / Il n'a rien dans l'esprit, il n'a rien dans le cœur ; / De ce qui l'environne il n'est qu'imitateur ; / Il répète les noms de devoir, de justice ; / Il agit en machine ; et c'est par sa nourrice / Qu'il est juif ou païen, fidèle ou musulman, / Vêtu d'un justaucorps, ou bien d'un doliman. » / Oui, de l'exemple en nous je sais quel est l'empire. / Il est des sentiments que l'habitude inspire. / Le langage, la mode et les opinions, / Tous les dehors de l'âme, et ses préventions, / Dans nos faibles esprits sont gravés par nos pères, / Du cachet des mortels impressions légères. / Mais les premiers ressorts sont faits d'une autre main : / Leur pouvoir est constant, leur principe est divin.' Although Voltaire agrees that our moral outlook is framed significantly by both language and education, he maintains that the fundamental principles of morality are universal and do not merely depend on convention.

hedonist conception of the Good leads him to a consequentialist notion of morality. Helvétius then employs this conception of morality as his fundamental position for his critique of contemporary society. In addition, Helvétius also follows his empiricist predecessors in their refusal to ascribe human understanding any innate ideas or properties, leading him to the conclusion that the content of the mind is entirely derived from sensation. Like Locke and Condillac, Helvétius then concludes that the moral outlook of individuals is therefore dependent upon both proper education and a correct understanding of the conventional meaning of moral terms. This conclusion then provides Helvétius with the theoretical framework for his proposed reforms of society.

At the same time, we have seen that Helvétius simplifies the theories of his empiricist predecessors, leading to a reductionist conception of human nature. Thus Helvétius presents man as an instinctual being, driven by desires and entirely determined by the ideas that happen to comprise his mind. Helvétius's simplified conception of human nature then also magnifies many positions in his moral and political philosophy. Thus in previous chapters we have seen that their hedonist conception of the Good had led Hobbes, Locke and Condillac to a consequentialist theory of morality. Yet it is only Helvétius who draws the unmitigated utilitarian conclusion from this premise that it is the purpose of morality and the state to cater to the desires and aversions of individuals. Furthermore, we have previously seen that following their empiricism, both Locke and Condillac considered it crucial that human beings are exposed to the right kind of ideas – hence their interest in the reform of both language and education. Yet following his theory of the mind, Helvétius takes these concerns to a whole new level. Thus for Helvétius, man can only become virtuous by being inculcated with moral ideas that direct him towards the common good. As Helvétius appears to suggest that it is ultimately the responsibility of the state to disseminate the right moral notions among its citizens, this position inadvertently leads him to what may be described as a totalitarian conclusion.

5. From Empiricism to Perfectibility: Rousseau on the Flexibility of Human Nature and Morality

Introduction

Like the thinkers discussed in our previous chapters, Rousseau bases his moral and political philosophy on a theory of human nature that describes man's intellectual abilities and explains the various motivations behind human behaviour.[434] Yet in many respects, Rousseau's theory of human nature exceeds that of his predecessors. To begin with, Rousseau clearly had profound insight into human psychology, providing explanations for the behaviour of individuals as well as groups of people compelling enough that they are still referenced in order to explain contemporary social and political developments.[435] Yet the most important innovation of Rousseau's theory of human nature compared to that of his predecessors is that he no longer viewed human nature as universal and largely fixed, but rather as fluid and subject to historical development largely determined by external factors. In Rousseau's own words, his predecessors had failed to appreciate that human nature is 'perfectible.'

Yet while this idea of the perfectibility of human nature is clearly an innovation, we will discover that Rousseau's theory of human nature is nonetheless elaborated on preconceptions largely adopted from Hobbes, Locke, and Condillac. Thus we will see that a crucial component of Rousseau's theory of the perfectibility of man is the idea, derived from Locke and above all Condillac, that reason is not innate to human nature, but rather an ability acquired gradually and dependent upon the development of language. Concurring with his empiricist predecessors that the composition of complex ideas requires the use of words, Rousseau also thinks that the human intellect has only developed beyond the level of that of other animals with the invention of language. Accordingly, Rousseau thinks that the simultaneous development of language and human understanding is the cause for the evolution of all other perfectible properties of human nature. Furthermore, we will demonstrate that Rousseau, influenced by

[434] Rousseau, *Second Discourse* Preface i p.124: 'The most useful and the least advanced of all human knowledge seems to me to be that of man, and I dare say that the inscription on the Temple at Deplhi alone contained a more important and more difficult Precept than all the big Books of the Moralists.' While the Temple at Deplhi was inscribed with 147 aphorisms, Rousseau is probably referring to its most famous precept: 'Know thyself.' Cf. preface v p.125-126.

[435] E.g. Pankaj Mishra, 'How Rousseau predicted Trump' *The New Yorker* (August 1, 2016).

Hobbes's account of practical deliberation, maintains that man's passions have expanded in proportion to the development of reason. Finally, we will see that Rousseau ascribes to human nature two properties not encountered in the philosophies of his predecessors with an important role in the constitution of morality, namely the sentiment of pity and the development of *amour propre* or self-esteem. Yet we will show that even these two properties are modified by the perfection of human understanding.

Our purpose in this chapter will be threefold. First of all, we will demonstrate the extent of Rousseau's indebtedness to preceding empiricist thinkers, and in particular to the philosophers discussed in our preceding chapters. We will then see that Rousseau not only borrowed both his methodology as well as many of his preconceptions from his predecessors, but also that Rousseau develops some of his theories from ideas already implicit in the philosophies of Hobbes and Condillac. Secondly, we will also discuss various ways in which Rousseau's conception of human nature provides the basis for his moral and political philosophies – thereby pursuing a line of inquiry also prominent in our previous chapters. Finally, as a result of this investigation of Rousseau's conception of human nature and its relation to his ethics, we will also be in the position to engage with scholarly interpretations of Rousseau moral and political philosophy. For instance, our investigation of Rousseau's moral epistemology will lead us to contest a reading of Rousseau's moral philosophy as a form of Platonism. Furthermore, our account of Rousseau's conception of human nature will also allow us to contextualise the political philosophy of the *Social Contract*, presenting it not merely as a theory of popular sovereignty, but rather as a plan for the revival of two properties of life in the state of nature that had accorded natural man a felicitous existence: liberty and equality. Finally, our interpretation will accept Rousseau's own contention that his thought forms a consistent unity.[436] As Rousseau's coherence is sometimes questioned by modern scholars, our own reading will also try to show that his philosophy may be interpreted without having to question this contention.[437]

---

[436] Rousseau, *Social Contract, with Geneva Manuscript and Political Economy* ed. by Roger Masters trans by Judith Masters (Boston 1978) II.iv n48 p.62; II.v p.65. James Delaney, *Rousseau and the Ethics of Virtue* (London 2006) 131-132 and Arthur Melzer, *The Natural Goodness of Man: On the System of Rousseau's Thought* (Chicago 1990) 4 accord similar weight to Rousseau's contention of coherence.

[437] Although, Ernst Cassirer already remarked in his essay *The question of Jean-Jacques Rousseau* ed. and trans. by Peter Gay (Bloomington 1954) p.53 that the interpretation of Rousseau as

Rousseau expounds his conception of human nature in two separate yet related accounts. On the one hand, Rousseau provides a theoretical description of the properties and abilities of human nature. Like previous empiricist philosophers, Rousseau thus provides an account of the origin of man's ideas, the motivations of his actions, and the development of the faculties of the mind. On the other hand, Rousseau also provides a historical account of the concurrent development of human nature and society. Following his conception of the perfectibility of man, Rousseau presents a conjectural history of the development of human nature following the invention of language, institution of society, and resulting expansion of human understanding. Our discussion of Rousseau's theory of human nature will begin with a concise overview of this latter conjectural history of the development of human nature. Subsequently, we will provide a more detailed exploration of Rousseau's conception of man by discussing his theoretical account of human nature. We will investigate Rousseau's epistemology, his theory of the passions and account of practical deliberation, as well as his ideas on how the feeling of *amour propre* and sentiment of pity influence man's behaviour. On the basis of our detailed investigation of Rousseau's conception of human nature, we will ultimately be in the position to provide a distinct perspective on his moral and political philosophy.

The State of Nature and the Perfectibility of Man

Following Hobbes, early-modern moral philosophers often commenced their discussion with a version of the thought experiment of the state of nature. In our first chapter, we have seen that Hobbes constructs his moral philosophy upon the premise that the state of nature is a state of war, caused by man's fear of death and desire for power to fulfil incessant appetite, leading to a scramble over resources and a conflict of rights. Hobbes then presents the erection of the sovereign as the only viable solution to these problems of life in the state of nature. While subsequent thinkers, such as Locke and Pufendorf, amend Hobbes's account on some points, they also conduct a thought experiment of the state of nature in order to exhibit why, in view of certain properties of human

incoherent is most common in antiquated readings. Nonetheless, a more recent examples of an interpretations that contest Rousseau's profession of coherence can be found in John Hall, *Rousseau: An Introduction to his Political Philosophy* (Cambridge Mass 1973) p.7; John Charvet, *The Social Problem in the Philosophy of Rousseau* (Cambridge 1974) 2-3.

nature, life in the state of nature is beset with difficulties, and to demonstrate that these may only be resolved by the institution of society and centralised authority.

Rousseau also accords the thought experiment of the state of nature a prominent place in his philosophy. But both the purpose and the conclusions of his account are fundamentally opposed to those of his predecessors. In fact, Rousseau's depiction of the state of nature is almost entirely the inverse of Hobbes's account. The fundamental premise of Rousseau's account is that human beings are naturally good, and that consequently the state of nature is almost entirely peaceful.[438] Rather than tormented by ever expanding desires, Rousseau's natural man is content with bare necessities. In addition, Rousseau contends that natural man does not have any notion of death, and consequently does not fear it.[439] Furthermore, Rousseau maintains that natural man is for the most part a solitary being that only encounters other human beings by chance.[440] Rousseau thinks that these occasional meetings would generally have a peaceful outcome, as natural man is moved by a sense of pity that has been almost entirely stifled in civilised man.[441] Elsewhere, Rousseau concedes that an encounter between humans in the state of nature may sometimes lead to a confrontation. But he thinks that in an abundance of resources, human beings in the state of nature generally coexist peacefully.[442] Accordingly, Rousseau's account of life in the state of nature neither calls for the institution of society nor of any form of centralised authority.

---

[438] Rousseau, *Lettre à Beaumont* in: *Œuvres complètes de Jean-Jacques Rousseau* ed. by Bernard Gagnebin (Paris 1969) IV p.935-936: 'Le principe fondamental de toute morale, sur lequel j'ai raisonné dans tous mes écrits, et que j'ai développé dans ce dernier avec toute la clarté dont j'étois capable, est que l'homme est un être naturellement bon, aimant la justice et l'ordre ; qu'il n'y a point de perversité originelle dans le cœur humain, et que les premiers mouvemens de la nature sont toujours droits.' Cf. Rousseau, *Émile* in: *Œuvres complètes* II p.322; *Second Discourse* note IX p.197-198.

[439] Rousseau, *Second Discourse* I.xix p.142.

[440] Rousseau, *Second Discourse* I.xxxiii p.149.

[441] Rousseau, *Second Discourse* preface x p.127.

[442] Rousseau *Essay on the Origins of Languages* in: *The Discourses and other political writings* ed. by Gourevitch IX.vi p.268-269: 'These times of barbarism were the golden age; not because men were united, but because they were separated [...] no one knew or desired anything but what was already to hand: his needs, far from drawing him closer to those like himself, draw him away from them. Men may have attacked one another upon meeting, but they rarely met. Everywhere the state of war prevailed, yet the whole world was at peace.'

The antithesis between Rousseau's and Hobbes's characterisations of life in the state of nature follows from a fundamental disagreement on the conceptualisation of human nature. To be sure, Rousseau's objection against Hobbes and other moral philosophers including Pufendorf and Locke is not that they have employed the wrong methodology in analysing human nature. Like Hobbes, Rousseau also employs introspection and observation of human behaviour as his primary methods. Neither does Rousseau deny that the conceptions of human nature of his predecessors aptly characterise the behaviour of modern humans. Rather, Rousseau's objection against his predecessors is that following their method of introspection, they had accepted the characteristics of their own mind and of others living in modern society as the original properties of human nature. For Rousseau, the primary error of Hobbes and his followers had been to simply assume that human nature is both universal and static. Yet Rousseau thinks that their efforts have only exhibited the characteristics of human nature as it is today, after it has been both perfected and disfigured by the concurrent development of reason and society. According to Rousseau, Hobbes and his followers therefore did not succeed in conceptualising human nature is it had been in the state of nature.[443]

Rousseau thus maintains that the shift from the state of nature to society not only involved the emergence of morality and the institution of political authority, it also caused a transformation of human nature itself. To begin with, Rousseau maintains that humans in the state of nature would have had an intellect equal to other animals of similar size. Furthermore, Rousseau postulates that natural man did not yet have the use of language. In addition, man in the state of nature also lacked any sense of morality, as moral concepts are inconceivable without the use of words.[444] Finally, Rousseau maintains that modern man's incessant desire for pleasure, power, and esteem develop only with the advent of society, and that natural man is content to satisfy only his most basic urges. Thus Rousseau maintains that Hobbes 'improperly included in Savage man's care for his preservation the need to satisfy a multitude of passions

[443] Rousseau, *État de guerre* in: in: *Œuvres complètes* III p.611-612: 'The error of Hobbes and the philosophers [is to have confounded] natural man with the men that they have before their eyes, and of transporting into one system a being that can subsist only in another.' Cf. *Émile* IV p.535. For Rousseau's introspective method and a comparison with Hobbes see Melzer, *The Natural Goodness of Man* 31 and Cassirer, *The Question of Jean-Jacques Rousseau* 51.

[444] Rousseau, *Second Discourse* I.xxxiv p.150

that are the product of Society and have made Laws necessary.'[445] Describing natural man as more similar to animals than to modern human beings, Rousseau maintains that previous moral philosophers including Hobbes have not identified the properties of natural man, and rather 'spoke of Savage Man and depicted Civil man.'[446] Accordingly, Rousseau maintains that human nature is flexible and has changed dramatically over the course of history. Rousseau himself uses the term 'perfectibility' to refer to the flexibility of human nature. This term may suggest that Rousseau considered human nature as proceeding from primitive origins towards a certain form of perfection. Yet Rousseau does not retain a teleological conception of human nature. He merely thinks that human beings possess a number of latent faculties that remain dormant in the state of nature, but which can be developed in a variety of ways depending on external circumstances.

Rousseau presents the perfectibility of human nature as the most significant property distinguishing man from the animals.[447] Yet Rousseau thinks that this perfectibility depends in turn on two properties unique to human nature. The first of these is human liberty. Rousseau's statements on the question whether humans possess freedom of the will are somewhat ambiguous, giving rise to scholarly disagreement on the issue. On the one hand, Rousseau sometimes appears to defend metaphysical arguments in favour of the idea that human beings have free will.[448] On the other hand, Rousseau elsewhere provides a largely mechanistic account of the process of the formation of the will, suggesting a compatibilist view on human liberty similar to that of Hobbes and Locke.[449] Yet we will presently forego a more extensive discussion of this issue, as

[445] Rousseau, *Second Discourse* I.xxxv p.151.

[446] Rousseau, *Second Discourse* exordium v p.132.

[447] Rousseau, *Second Discourse* I.xvii p.141: 'there is another very specific property that distinguishes between them [man and animals], and about which there can be no argument, namely the faculty of perfecting oneself; a faculty which, with the aid of circumstances, successively develops all the others, and resides in us, in the species as well as in the individual.'

[448] Rousseau, *Émile* in: *Œuvres complètes* IV p.586. Maurice Cranston, 'Rousseau's Theory of Liberty', in : Robert Wokler (ed.), *Rousseau and Liberty* (Manchester 1995) 231-243, 231 therefore ascribes to Rousseau the view that humans do possess freedom of the will.

[449] John Duncan, 'Perfectibility, Chance, and the Mechanism of Desire Multiplication in Rousseau's *Discourse on Inequality*' in: Mark Blackwell et al. (eds.), *Rousseau and Desire* (Toronto 2009) 17-45, p.21 argues that Rousseau does not ascribe to man the capacity for free agency, pointing to the sometimes mechanistic metaphors Rousseau employs to describe human thought processes. In addition, Duncan maintains that man's perfectibility excludes the possibility that man has liberty of will.

the precise interpretation of Rousseau's conception of human liberty is not of crucial importance to our wider discussion.[450] For our present purposes, it is sufficient to understand Rousseau's conception of human liberty merely as the capacity of humans to adapt their way of life. Thus whereas the behaviour of animals is strictly regulated by instinct, both history and what today we might call anthropological evidence have shown that human beings are free to choose many different ways of living. Rousseau thereby presents man's independence from instinct as empirical evidence for human liberty.[451] At the same time, this independence from instinct is also an important precondition for the perfectibility of human nature. Rousseau thereby suggests that man's perfectibility is only made possible by man's liberty.

The second property of human nature Rousseau considers essential to man's perfectibility is the ability – itself dormant in natural man – to acquire conventional language.[452] According to Rousseau, language is not merely a means for communication essential to sociability, but it is also essential to the development of all of man's other perfectible qualities. The main reason why Rousseau maintains that language enables man's perfectibility is that, influenced by the empiricist philosophers we have discussed in previous chapters, he thinks that words are essential to the composition of complex ideas. This ability, unique to human beings, is in turn fundamental to all other of man's perfectible properties. Among the complex ideas that human beings are able to construct through the use of words are abstract and universal ideas, as well as moral concepts. Consequently, Rousseau thinks that all of man's perfectible qualities, including the capacity for intellectual reason, foresight, and the ability to live sociably according to shared moral conventions, are ultimately the consequence of the ability to construct complex ideas, which is in turn made possible by the use of words.

In the following, we will investigate Rousseau's theory of the perfectibility of human nature in greater detail. Rousseau provides two separate accounts of the theory of human perfectibility. His first explanation, found in his *Second Discourse*, is a conjectural history of human development, detailing how through successive innovations natural man gradually evolved into the civilised

[450] Charvet, *The Social Problem in the Philosophy of Rousseau* 10 agrees with this assessment.

[451] Rousseau, *Second Discourse* I.xvi p.140-141.

[452] Rousseau, *Essay on the Origin of Languages* I.xiv p.252 recognises that while some animals also have means of communication, unlike humans they do not use conventional signs as language.

man of modern times. Yet in his educational treatise *Émile*, as well as scattered throughout the conjectural history of the *Second Discourse*, Rousseau also provides a more theoretical account of the perfectible properties of human nature.[453] Our investigation will focus mainly on this theoretical description of human nature. Yet as both accounts are intertwined, we first require an outline of Rousseau's conjectural history in order to fully understand his theory of human nature. We will therefore begin our discussion with a brief overview of Rousseau's account of the state of nature and development of society, and afterwards turn our attention towards his theoretical account of human nature.

From the State of Nature to Society

Rousseau's conjectural history subdivides human development into four different stages.[454] As we have seen above, Rousseau thinks that in the state of nature man has a solitary and animal existence. This first stage is followed by a second era of slowly developing sociability. Rousseau recounts how natural man must have been forced to abandon his innocent and happy existence due to a shortage of resources following natural disasters and population growth.[455] In the period that followed, human beings slowly settled into what might be called a tribal existence. Living in small bands the size of extended families, these tribal humans gradually learned to live together.[456] The most important factor in this early development of sociability was the emergence of primitive language. It not only allowed tribal humans to cooperate, but also to form lasting and more extensive social and moral bonds. At the same time, primitive language also expanded man's dormant intellectual abilities, as it allowed tribal man to compose ever more complex ideas, thereby gradually causing the development of reason.[457] Rousseau thinks that most primitive tribes found in the modern world have attained this tribal stage of existence. Rousseau also considers this tribal stage the happiest period in human existence, during which man progressed

[453] Rousseau, *Émile* I p.251 states that while ostensibly an educational treatise, the *Émile* also aims to establish a conception of human nature.

[454] It is conjectural because Rousseau begins his account by stating at *Second Discourse* exordium 6 p.132: 'The Inquiries that may be pursued regarding this Subject ought not to be taken for historical truths, but only for hypothetical and conditional reasonings.'

[455] Rousseau, *Essay on the Origin of Languages* I.ix.27 p.274.

[456] Rousseau, *Second Discourse* II.xii p.164.

[457] Rousseau, *Second Discourse* I.xxx-xxxi p.148-149.

beyond a purely animal existence, while the conflicts and oppression that characterise subsequent eras are still largely absent.[458]

Rousseau locates the conflict and disorder that characterises previous descriptions of the state of nature, such as those of Hobbes, Locke and Pufendorf, only in the third era of his conjectural history of human development. During this period, man's enhanced intellect enabled the invention of agriculture and metallurgy, prompting man to switch from a nomadic to a sedentary lifestyle. Yet the most important novelty of this period was a moral invention: the notion of property. Although not made with malicious intent, the invention of private property is in Rousseau's view the principal origin of man's subsequent hardships.[459] In the period before the invention of property, humans had lived in equality, as both land and natural resources were considered common goods. Consequently, men had no reason to envy or desire each other's possessions. Yet when some men converted to a sedentary lifestyle and started to consider their land and its yields as their own, they unwittingly instituted a form of inequality. This not only caused humans to envy and contest each other's private possessions, but also required the propertied to actively defend their assets. As a result, Rousseau thinks that only during this third stage in human development the state of conflict arose that Hobbes locates in the state of nature.[460]

---

[458] Rousseau, *Second Discourse* II.xviii p.167.

[459] Rousseau, *Second Discourse* II.i p.161. Interestingly, Rousseau appropriates one of the 'moral demonstrations' we have discussed in our chapter on Locke to support this argument. Thus we have seen that Locke tried to show the possibility of a demonstrative moral science by presenting two examples of moral demonstration. Among these is the supposed demonstration that without property there can be no injustice. Rousseau takes this contention out of context, ending his discussion of the invention of property and its consequences in the *Second Discourse* II.xvii p.166 by stating that: 'according to the axiom of the wise Locke, "*where there is no property, there can be no injury.*"'

[460] Rousseau, *Second Discourse* II.xix-xxvi p.167-170. Some modern anthropologists, palaeontologists and archaeologists have arrived at similar conclusions. Thus in his book *Sapiens, A Brief History of Humankind* (Vintage: London 2011) 87-110, Yuval Noah Harari maintains that nomadic hunter-gatherers generally had higher standards of living and a longer life expectancy than humans living after the agricultural revolution – at least until the beginning of the twentieth century. Furthermore, a recent paper by Timothy Kohler et al., 'Greater post-Neolithic wealth disparities in Eurasia than in North America and Mesoamerica' *Nature* 551 (2017) 619-622 demonstrates that inequality increased after the agricultural revolution. In addition, Stephen Shennan, 'Regional population collapse followed initial agricultural booms in mid-holocene Europe' *Nature Communications* 4 (2013) suggests that with the invention of agriculture, food production and life expectancy became less rather than more secure.

Rousseau finally locates the institution of civil society and the state in the fourth era of his conjectural history of human development. Yet Rousseau does not follow Hobbes in portraying the institution of the state as the outcome of a voluntary social contract in which citizens subject to a sovereign in order to end the state of war detrimental to the interests of all. Rather, Rousseau thinks the state was created to protect the interests and property of the wealthy and privileged, thereby institutionalising the inequality that had developed during the third era of human development.[461] Whereas previously there had only been natural inequality between human beings caused by differences in physical strength and intellectual capacity, the emergence of civil society introduced a moral inequality far more consequential. Furthermore, the maintenance of this inequality required the institution of a government that safeguarded the interests of the privileged rather than the common good of society as a whole. As a result, Rousseau thinks that as they have arisen historically, most modern states – including contemporary France – have developed into despotic societies.[462]

Rousseau's conjectural history therefore does not portray the gradual emergence of society as a positive development. Yet by themselves, inequality and despotism are not even its most regrettable consequences. Even more lamentable is the fact that in concurrence with the emergence of society, human nature itself has been modified in such a way that modern man has become a contradictory and almost inevitably miserable being. Thus while man's intellectual awakening has not only allowed him to acquire scientific knowledge and devise all kinds of ingenious inventions, it has also caused an inflation of his desires. Whereas natural man had been content with the satisfaction of his primary needs, civil man's increased imagination and awareness of the incalculable forms of pleasure available in modern society has inflamed his passions far beyond his natural desires. As it is hardly ever possible to satisfy all these artificial passions, Rousseau thinks that civil man is doomed always to remain unhappy, as certain forms of luxury or achievement are bound to remain ever beyond his grasp.

---

[461] Rousseau, *Second Discourse* II.xxiv p.169; II.xxx p.172.

[462] Rousseau, *Second Discourse* II.xxxii p.173: 'All ran toward their chains in the belief they were securing their freedom; for while they had enough reason to sense the advantages of a political establishment, they had not enough experience to foresee its dangers; those most capable of anticipating the abuses were precisely those who counted on profiting from them.'

Yet the inflation of the passions of civilised man is not the only source of his misery. In addition, Rousseau thinks that by living together in society, human beings develop the urge to be esteemed and admired by their fellows. Rousseau maintains that human beings are naturally endowed with what he refers to as *amour de soi* – a benign form of self-love that induces man to care for his own preservation. Yet during the emergence of society, this self-love generally mutates into a form of self-esteem which Rousseau calls *amour-propre*. After natural man's intellectual awakening, human beings become able to compare themselves to the people with whom they now live together. Consequently, the harmless *amour de soi* of natural man develops into *amour-propre* that causes civilised man to value himself over others. This self-esteem then causes civilised man to long for a confirmation of his significance in the form of the esteem of his fellows. *Amour-propre* is therefore the main reason for civilised man's desire for honour and reputation, and thereby the primary cause of ambition. Yet Rousseau thinks that *amour-propre* is also an important cause of civilised man's misery. As it induces us to value ourselves more than others, the demands of *amour-propre* can be met only if others also esteem us more than they value themselves. But because everyone in modern society is subject to the same form of selfishness, no individual will be able to obtain the desired esteem.

Rousseau thus presents man's development as the unfortunate progression from a simple yet innocent creature, content within its environment, into an intellectually more advanced yet covetous being in need of restraint by an oppressive sovereign. The perfectibility of human nature has therefore contributed more to the misery than the happiness of man.[463] For this reason, Rousseau's description has been interpreted by some scholars as a secular version of Christian theodicy.[464] Like Adam and Eve in the Garden of Eden,

[463] Rousseau, *Second Discourse* I.xvii p.141: 'It would be sad for us to be forced to agree that this distinctive and almost unlimited faculty, is the source of all of man's miseries; that it is the faculty which, by the dint of time, draws him out of that original condition in which he would spend tranquil and innocent days; that it is the faculty which, over the centuries, causing his enlightenment and his errors, his vices and his virtues to bloom, eventually makes him his own and Nature's tyrant.'

[464] This is the general argument of John Scott in his paper: 'The Theodicy of the Second Discourse: The Pure State of Nature and Rousseau's Political Thought' *The American Political Science Review* 86 (1992) 696-711. Cassirer, *The Question of Jean-Jacques Rousseau* 76 maintains that Rousseau's revision of theodicy made society rather than God responsible for the faults of man. According to Cassirer, this revision is at the root of all contemporary political debates, as the modern state has become responsible for the well-being of its citizens.

natural man has pursued things that may have seemed attractive or profitable at first, but which for various reasons turned out to make him decidedly unhappy – losing his innocence in the process. Yet unlike in the case of Christian theodicy, Rousseau's account of the fall of man has been unknown until he uncovered it. The reason why others have overlooked the development from natural man into civilised man and considered human nature as fixed, is not merely the fact that the changes to man's nature took place very gradually and over a long period of time. It is also because the nature of civilised man has been transformed to such an extent that it no longer exhibits any traces of natural man. Furthermore, modern man has been oblivious to the fact that civil society and the despotic state necessarily make him unhappy, as his predicament is legitimised by deeply ingrained moral conventions – or prejudices – acquired from childhood. In Rousseau's view, it required a social and intellectual outcast like himself to see through these prejudices and beyond man's artificial properties to uncover man's true nature.[465]

Rousseau's conjectural history thus presents the transformation of natural into civil man as a predominantly negative development. To fully understand why Rousseau was so pessimistic about life in society, we will now proceed with a more comprehensive discussion of his theoretical account of human nature. This discussion will explain in greater detail why Rousseau thinks that the development of man's faculties has had such unfortunate consequences. To be sure, Rousseau is not a primitivist, as is sometimes thought, who advocates a return to life in the state of nature.[466] Rather, in his theoretical account of man's perfectibility, Rousseau identifies several ways in which the development of human nature has backfired. While the development of man's faculties dramatically increases his capabilities, it may also configure human nature in such a way that life in society will make human beings both unhappy and ignorant. And Rousseau maintains that unfortunately, the historical development of human nature has indeed ensured the misery of civilised man. In the

[465] Rousseau, *Émile* II p.348: 'Lecteurs, souvenez-vous toujours que celui qui vous parle n'est ni un savant ni un philosophe; mais un homme simple, ami de la vérité, sans parti, sans sistême; un solitaire qui vivant peu avec les hommes a moins d'occasions de s'imboire de leurs préjugés, et plus de tems pour réflechir sur ce qui la frape quand il commerce avec eux.' Judith Shklar, *Men and Citizens: A Study of Rousseau's Social Theory* (Cambridge 1969) p.46 speculates that the figure of natural man as happy and indolent may be inspired by Rousseau's own experience as a vagabond.

[466] Rousseau, *Second Discourse* note IX.xiv p.203 states that a return to the state of nature is neither realistic nor desirable.

following, we will thus not only explore Rousseau's theoretical account of human perfectibility, but also detail how Rousseau thinks that the development of man's faculties has gone awry. Finally, we will also discuss some of the ways in which Rousseau thinks this situation may be remedied.

Ideas, Judgment, and Reason

Rousseau's theoretical explanation of man's perfectibility can be subdivided into three distinct accounts that each addresses a certain component of human nature. Subsequently, we will deal with Rousseau's theory of practical deliberation as well as with his account of how *amour propre* and the sentiment of pity affect moral relations between humans. But we will begin by discussing his theory of the development of human understanding. Rousseau was by no means an original thinker on epistemology. Neither did he take the time to explain his theory of human understanding in any systematic way. Nonetheless, Rousseau's interest in epistemology, particularly in relation to morality, has often been overlooked by commentators.[467] In fact, one of the primary arguments of this chapter is that even if Rousseau was hardly interested in epistemology for its own sake, his theory of human nature, and by extension his moral philosophy, are nonetheless dependent upon a theory of human understanding derived from earlier empiricist philosophers.

In his ideas on epistemology, Rousseau often relied on the empiricism of Locke's *Essay*, which he had closely studied from his early education onwards.[468] Furthermore, as Rousseau himself confesses, he was also heavily influenced by Condillac, with whom he had been friends during his years in Paris.[469] While Rousseau embraces some of Condillac's innovations, he is also highly critical of the reductionist epistemology Helvétius developed on the basis of Condillac's

[467] Ryan Patrick Hanley, 'Rousseau's Virtue Epistemology' *Journal of the History of Philosophy* 50 (2012) 239-263, p.239 agrees with this assessment.

[468] Peter Jimack, *La Genèse et la rédaction de l'Emile de J.-J.Rousseau* (Geneva 1960) p.288. John Scott, 'Rousseau's Unease with Locke's Uneasiness' in: Eve Grace and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge 2013) 295-311, 303 maintains that the broad outlines of Rousseau's epistemology are 'clearly Lockean in inspiration.'

[469] Leo Damrosch, *Jean-Jacques Rousseau: Restless Genius* (Boston 2005) 152-154. Duncan, 'Perfectibility, Chance, and Desire Multiplication' 30-33 also places Rousseau in the context of Condillac and Lockean Empiricism, but apart from pointing out the personal connections between Rousseau and Condillac, does not provide any details on the role of Condillac's philosophy as a source of inspiration to Rousseau.

theory.[470] In response, Rousseau often remains closer to Locke's positions.[471] Yet while Rousseau borrows many of his epistemological preconceptions from his predecessors, we will see that a number of different conclusions lead him to an altogether different conception of human nature. In the following sections, we will first review Rousseau's statements on human understanding, and subsequently discuss his views on the role of language in cognition. When we have acquired a solid grasp of his epistemology, we will be able to fully comprehend why Rousseau thought that the expansion of man's cognitive faculties has not been a predominantly positive development.

Like any true empiricist, Rousseau takes the position that the mind is a *tabula rasa* at birth and that all ideas are acquired through the senses as his fundamental axiom.[472] Rousseau thereby follows Condillac in discarding Locke's notion that the mind may also acquire ideas of its own operations through a form of introspection called reflection. Like Condillac, Rousseau thus employs the term reflection to refer loosely to any form of reasoning.[473] Furthermore, Rousseau adopts the distinction, introduced by Locke, between simple ideas of sensation and complex ideas created by the mind by compounding multiple ideas. Yet Rousseau does not take on Locke's more subtle distinctions between different kinds of ideas. As he uses the terms complex or intellectual idea interchangeably to refer to any form of compounded idea, it is not always easy to determine precisely to which kind of ideas he is referring. It is clear though that moral concepts count as intellectual ideas, while Rousseau likely also considers what Locke would call abstract or universal ideas of substances as intellectual ideas.

Another position that Rousseau seems to have borrowed from Locke is that ideas are the 'Workmanship of the Mind.'[474] Thus while it is clear that the content of our ideas is ultimately derived from experience, Rousseau thinks that

---

[470] Rousseau's opposition to Helvétius' ideas is evident from the marginalia found in his personal copy of *De l'esprit,* which are reproduced in *Œuvres complètes* IV. Jimack, *La Genèse et la rédaction de l'Emile* p.319 speculates that the *Confession of the Savoyard Vicar* found in the *Émile* was partly intended as response to Helvétius' philosophy. Although not mistaken, it is probably more accurate to state that the *Confession* was a response to positions held by a number of contemporary French enlightenment thinkers, which include monist materialism, determinism and the notion of morality as a prudential rule indicating how to satisfy desires and avoid pain.

[471] Shklar, *Men and Citizens* p.37 and Scott, 'Rousseau's Unease with Locke's Uneasiness', 296 agree with this assessment.

[472] Rousseau, *Lettres Morales* III in: *Œuvres complètes* IV p.1092; *Émile* II p.370.

[473] Rousseau, *Essay on the Origin of Languages* IX.iii p.268.

[474] Locke, *Essay* III.v.4 p.430.

our ideas are not identical to sensations. Rather, Rousseau maintains that sensations merely provide the raw material out which the mind subsequently constructs its ideas:

> Dans la sensation, le jugement est purement passif, il affirme qu'on sent ce qu'on sent. Dans la perception ou idée, le jugement est actif ; il rapproche, il compare, il détermine des rapports que le sens ne détermine pas. Voila toute la différence, mais elle est grande. Jamais la nature ne nous trompe ; c'est toujours nous qui nous trompons.[475]

According to Rousseau, ideas are thus not just memorised sense impressions – or 'decaying sense' in Hobbes's words. Ideas are formed when mind analyses sensations and assembles them into ideas, meanwhile establishing connections to ideas it has previously compounded. The creation of ideas thus requires active judgment. Rousseau thereby explicitly contests Helvétius's views that the mind itself is nothing more than a collection of ideas, that all operations of the mind can be reduced to physical sensitivity, and that what is often called judgment is in fact nothing but sensing.[476] According to Rousseau, human beings are only able to proceed beyond the level of mere sensation because they possess a faculty of judgment that allows them to collect their experiences into ideas.[477]

In this respect, Rousseau again appears to follow Locke, who had also maintained that complex ideas are compounded at will, thereby tacitly assuming that the mind is an entity separate from its ideas and contains a faculty of judgment beyond its ability for sensation.[478] Yet unlike Locke, Rousseau does not assume that this faculty is innate to human understanding. Likely inspire by Condillac, Rousseau thinks that man's cognitive abilities are habits that develop

---

[475] Rousseau, *Émile* III p.481; cf. II p.344; IV p.571-573.

[476] Rousseau, *Notes sur "De l'esprit"* in: *Œuvres* IV p.1122-23. For an extended comparison between Rousseau's position and that of Helvétius see Jimack, *La Genèse et la rédaction de l'Emile* 320-326 ; Terence Marshall, 'Epistemology and Political Perception in the Case of Rousseau', in: Eve Grace and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge 2013) 76-122, 79-91.

[477] Marshall, 'Epistemology and Political Perception' 94 suggests that Rousseau's theory that judgment is an operation of the mind distinct from mere sensation is 'anticipating Kant's "Copernican Revolution" in epistemology.' Yet our own reading suggests that we should rather consider Rousseau's theory as a more explicit formulation of Locke's tacit assumption that the mind is an active entity distinct from its ideas.

[478] This idea that man has a separate mental faculty and is free to construct his own ideas may be related to Rousseau's view that unlike animals, who are ruled by instinct, human beings have the freedom to diverge from their original constitution.

only gradually and require the proper education to be perfected.[479] Thus Rousseau repeats Condillac's thought-experiment of the statue endowed with sensibility, inquiring what would happen if we would create a full-grown man with the mind of new-born infant. Rousseau's answer is not merely that the mind of this man would be wholly devoid of ideas. He would also lack the judgment required to analyse his sensations and catalogue them into ideas. In fact, Rousseau maintains that at first, this infant man would not even be able to make sense of experience, as he lacks the capacity to distinguish one object from another.[480]

As neither of these basic cognitive abilities is innate to human understanding, Rousseau follows Condillac in thinking that the ability to analyse sensation and compose ideas with the help of judgment develop only gradually.[481] To some extent, these abilities are acquired through experience itself. Thus Rousseau maintains that infants do not yet have the ability to reason, and that consequently their memory is filled with images rather than ideas.[482] Rousseau subsequently outlines the lengthy process by which children learn to distinguish one object from another. He also thinks that at some point children will instinctively start to connect the ideas of objects that occur successively, thereby leading to the ability to perceive causal relations through induction.[483] As human beings acquire by experience the ability to analyse sensations as well as at least some aptitude in judgment, they will also learn to interact with their environment. This is the reason why natural man can take care of his preservation, despite the fact that without any education or language, his intellectual development is entirely due to experience.

Rousseau maintains that by attaining greater aptitude in judgment – in other words, by becoming more skilled in compounding and relating ideas – human beings will increase their ability to reason. Like the empiricist thinkers we encountered in previous chapters, Rousseau therefore does not consider

[479] There may be a tension between the idea that man's mental operations are nothing but habits, and the notion that man possesses a faculty of judgment separate from the ideas acquired through sensation – a tension that is left unresolved in Rousseau's own writings.

[480] Rousseau, *Émile* I p.280. Rousseau thereby implicitly takes sides in the debate over the Molyneux question, as he claims that the mind does not have the innate ability to distinguish the objects of sensation, but requires practice to acquire it.

[481] Rousseau, *Émile* I p.247.

[482] Rousseau, *Émile* II p.344.

[483] Rousseau, *Émile* III p.482.

reason as an innate and autonomous faculty that may provide insight into eternal truths independent from experience.[484] Rather, Rousseau follows his empiricist predecessors in describing reason as nothing but the composite of all other faculties of the mind.[485] Reason is therefore not itself a faculty, but rather a term to indicate aptitude in analysing, compounding, and relating ideas. Accordingly, Rousseau maintains that: 'sitôt que l'on compare une sensation à un autre on raisonne. L'art de juger et l'art de raisoner sont exactement le même.'[486] In this sense, natural man and even animals can be said to possess some form of reason, as both are capable of comparing sensations. For Rousseau, the property that distinguishes man from the animals is thus not reason, but rather perfectibility, which in turn depends on the ability to construct complex ideas.[487] Indeed, Rousseau maintains that with regard to the intellect, the difference between man and animals is one of degree.[488] Accordingly, Rousseau uses the term 'intellectual reason' to refer specifically to man's acquired ability to reason with complex ideas.[489] Reason is thus not a faculty innate to human nature, but rather an aptitude in the comparison of ideas – a skill perfected only with much exercise.[490] Yet in order to proceed beyond the intelligence of the beast, natural man required an invention that serves both as a method of communication and a means to expand the scope of reason by enabling the composition of complex ideas. This invention was language.

Language and Ideas

As Rousseau thinks that the more advanced development of the faculty of reason requires the use of language, he thereby adopts an idea that we have encountered in all our previous chapters.[491] In fact, Rousseau himself admits that his ideas on

[484] Rousseau, *Lettre a C. de Beaumont* in: *Œuvres* IV p.951.
[485] Rousseau, *Émile* II p.317.
[486] Rousseau, *Émile* III p.486.
[487] Rousseau, *Second Discourse* I.xxx p.148.
[488] Rousseau, *Second Discourse* I.xvi p.140-141.
[489] Rousseau, *Emile* II p.417.
[490] Rousseau, *Lettre a C. de Beaumont* in ; *Œuvres* IV p.951 : 'Vous supposez, ainsi que ceux traitent de ces matieres, que l'homme apporte avec lui sa raison toute formée, et qu'il ne s'agit que de la mettre en œuvre. Or cela n'est pas vrai ; car l'une des acquisitions de l'homme, et même des plus lentes, est la raison.'
[491] The important role of language in Rousseau's theory of the development of human understanding is not always sufficiently recognised by modern scholars. Thus Duncan, 'Perfectibility, Chance, and Desire Multiplication' 36-37 discusses how according to Rousseau

this topic have been significantly influenced by the philosophy of Condillac.[492] Like Condillac, Rousseau views language as a uniquely human ability that is entirely responsible for man's perfectibility and development beyond a purely animal state.[493] Rousseau mentions two ways in which language enhances human cognition. First of all, Rousseau thinks that the structure of language supports human understanding, as he states that 'Grammar exercises and facilitates the operations of the Mind.'[494] Rousseau thereby appears to adopt a version of Condillac's view that language allows the mind to extend and consolidate the connections between ideas, producing a cognitive network much greater than could have been created merely through experience. Unfortunately though, Rousseau's statements on this issue remain somewhat cursory, making it difficult to establish whether he fully adopted this element of Condillac's theory.

Secondly, Rousseau adopts a view common to all the philosophers we have previously considered – namely the theory that at least to some extent, ideas depend on words. Again, Rousseau's view appears to be closest to that of Condillac, as he states that the mind can only acquire or understand more complex ideas with the help of words:

> Besides, general ideas can enter the Mind only with the help of words, and the understanding grasps them only by the means of propositions. That is one of the reasons why animals could not form such ideas, nor ever acquire the perfectibility that depends on them.[495]

Rousseau maintains that the uniquely human ability to understand and construct general ideas, which are not acquired directly from experience but rather constructed by the mind at will, depends entirely upon the use of words. In passing, Rousseau thereby affirms that this ability to construct complex ideas is the reason why human nature is perfectible. For Rousseau, these kinds of ideas are thus so dependent upon words that they cannot be conceived without them. Rousseau therefore does not think, as Locke had done, that words merely aid the

---

humans are able to acquire what he refers to as "incremental knowledge" which exceeds the mere prudence shared with animals, but entirely misses cognitive role of language in the development of this knowledge.

[492] Rousseau, *Second Discourse* I.xv p.145.

[493] Rousseau, *Essay on the Origin of Languages* I.xiv p.252.

[494] Rousseau, *Second Discourse* I.xxiv p.144.

[495] Rousseau, *Second Discourse* I.xxx p.148.

mind by functioning as headings or knots tying simple ideas together to form a complex idea. Thus we have seen previously that Locke held a conceptualist position, which entailed that while the mind is certainly aided by the use of words when compounding and considering complex ideas, it was not absolutely dependent upon them and in theory could do the same without language.

By contrast, Rousseau maintains in the passage quoted above that when considering complex ideas, 'the understanding grasps them only by the means of propositions.' The content of general ideas is therefore expressed in a definition which is itself again comprised of words. Rousseau therefore appears to adopt the nominalist position of Hobbes, which considers general ideas as purely linguistic entities whose meaning is circumscribed by a definition. This conclusion is further supported by Rousseau's statement that:

> Every general idea is purely intellectual; if the imagination is involved, the idea immediately becomes particular. Try to outline the image of a tree in general to yourself, you will never succeed.[496]

Although Rousseau's terminology is again somewhat vague, he clearly defends the nominalist position we have also encountered with Hobbes. This view entails that a general idea is in fact nothing but a word with certain reference circumscribed by its definition, and that when the mind attempts to imagine this idea, it will always picture a particular object within the frame of reference of the general term. Consequently Rousseau states that abstract ideas are 'conceived only by means of discourse.'[497] Rousseau's statement that 'every general idea is purely intellectual' therefore refers to the fact that general ideas depend upon a definition and are thus necessarily linguistic in nature.

The Development of Language and the Perfection of Human Nature

Rousseau's account of the interaction of language and ideas thus closely resembles that of previous thinkers, including Hobbes and Condillac. Yet by itself, this theory is not sufficient to explain why Rousseau claims that human nature is perfectible. For this reason, we must discuss a notion that we have also encountered in our discussion of Condillac: the idea that language is itself a

[496] Rousseau, *Second Discourse* I.xxx p.148.
[497] Rousseau, *Second Discourse* I.xxx p.148.

human invention that developed gradually over the course of history. We have seen in a previous chapter that Condillac had maintained that in view of the dependence of ideas on words, the fact that language is a historical artefact entails that human knowledge must have had a history as well. The first component of Rousseau's theory of human perfectibility could be considered an elaboration of this insight. Following Condillac's suggestion, Rousseau then provides a more extensive account of how the historical development of language has gradually expanded human understanding and changed human nature

This history of human cognition begins with the figure of natural man, who lives before the invention of language and must therefore be solitary and stupid. Rousseau's characterisation of natural man as having an animal existence may appear curious at first. But Rousseau's description of natural man is in some ways anticipated in the philosophies of both Hobbes and Condillac. We have seen that in their discussions of human understanding, both thinkers explained that as the ability to form complex ideas depends entirely on the use of words, so do foresight and the faculty of reason. As we have seen in a previous chapter, the epistemological consequences of this theory are explored at length by Condillac in his thought experiment of the sensitive statue. As its senses are enabled, the statue will start to seek out sensations of pleasure and avoid pain. But lacking the use of words, the statue is limited to a purely animal existence in which it merely reacts to the external stimuli of sensation and the feelings of pleasure and pain. In many ways, Condillac's statue-man serves as the model for Rousseau's natural man, who is likewise limited to considering particular ideas and consequently has very limited reason and foresight.[498]

Rousseau thus arrives at his figure of natural man by combining the theory that the development of human understanding requires language with the view that language itself is the product of a historical development. Rousseau's conclusion that at the dawn of man, human beings would have been comparable to other animals then follows from his realisation that because language is an artefact devised by humans themselves, there must have been a time before the invention of language. At this time humans had neither the ability to communicate nor to employ words to construct complex ideas. This is then the reason why Rousseau begins his conjectural history of human development with

[498] Rousseau, *Second Discourse* II.iii p.161. Hulliung, *The Autocritique of Enlightenment* 67 also makes this suggestion. Rousseau was clearly acquainted with Condillac's thought experiment, as he references it in his *Lettres Morales* III p. 1096.

a description of man as naturally stupid and solitary. The theory that human beings require language to reason, which had previously been confined to discussions on epistemology and philosophy of language, is thus adopted by Rousseau as the preconception to his own revolutionary account of the development of human nature and the concurrent emergence of society and morality. In short, whereas empiricist philosophers had previously described the mind as a *tabula rasa* at birth, Rousseau maintains that human nature must also have been a blank slate at the dawn of mankind.

Rousseau thinks that at some point natural man has been induced by external circumstances to leave his solitary and inane existence to form primitive societies. Consequently, nascent man also became subject to passions beyond the necessities of mere subsistence. Incited by the urge to satisfy these passions, nascent man devised a rudimentary form of language as a means to communicate his feelings and intentions to his fellows.[499] Yet Rousseau admits that he cannot find an adequate explanation of how natural man would have been able to acquire the use of words without either education or example. Rousseau identifies two conundrums, which he presents as objections to Condillac's account of the history of language, but which, as he himself admits, equally apply to his own theory. The first problem is that as a necessarily communal institution, language can only emerge among a group of people, while sociability – or the ability to live in a group – is itself only made possible by the use of language. The question is thus how natural man would have been able to devise a language while still solitary, and alternatively how natural man would have been able to be social without the use of language.[500] Rousseau's second problem takes a similar form. As the invention of language is clearly an intellectual achievement of the highest order, the question arises how it could have been accomplished by a creature whose intellectual abilities are still in its infancy. As natural man is entirely without the use of words, he also lacks the

[499] Rousseau, *Essay on the Origin of Languages* II.iii p.253. By emphasising that language emerged due to the urge to satisfy passions, Rousseau diverges consciously from thinkers like Condillac, who had proposed that language was first invented by primitive man in order to acquire the necessities of existence. Yet, as we have seen, Rousseau maintains that natural man had required neither language nor sociability to ensure his self-preservation. Consequently, Rousseau maintains that language only became necessary once nascent man became sociable and developed artificial passions.

[500] Rousseau, *Second Discourse* I.xxv p.145.

ability to construct complex ideas – the very ability he would require to invent language.[501]

Rousseau himself concedes that he does not have adequate solutions to these two problems, stating rather that 'by means which I cannot conceive, our new Grammarians started to expand their ideas and generalise their words.'[502] In addition, Rousseau thinks that the invention of language and simultaneous development of human understanding must have taken innumerable centuries to complete. Rousseau's subsequent account of the development of language again closely follows that of Condillac. Rousseau describes how language emerged slowly as primitive humans began to live in small social groups and started to use cries and gestures to indicate their feelings. Gradually, these exclamations developed into conventional signs that could signify ideas. During this evolution, which again took innumerable centuries to complete, language became more precise, as words became differentiated by their grammatical function into nouns, verbs, pronouns, etc. Yet Rousseau thinks this progress also diminished the ability of language to convey the feelings that had been indicated by the cries and gestures of primitive man.[503] This is the reason why rhetoric in modern languages such as French will never have the same rousing effect as the words of ancient orators speaking in Latin or Greek.[504]

Yet while the evolution of language may have diminished some of its ability to convey feeling, it also increased the ability of words to signify ideas. Thus with the help of words, nascent man was now able to compose ever more complex ideas. This ability to compose complex ideas then gave humans two new abilities. First of all, words allowed humans to construct complex ideas of

[501] Rousseau, *Second Discourse* I.xxvi p.146: 'New difficulty, even worse than the preceding one; for is Men needed speech in order to learn how to think, they needed even more to know how to think in order to find the art of speech; and even if it were understood how the sounds of the voice came to be taken for the conventional interpreters of our ideas, it would still leave open the question of what could have been the interpreters of that convention for ideas, which, having no sensible object, could not be pointed to by gesture or by voice.' Cf. Lifschitz, *Language and Enlightenment* 35-37 for a discussion of this objection in the context of French Enlightenment debates on language.

[502] Rousseau, *Second Discourse* I.xxxi p.148.

[503] Rousseau, *Essay on the Origin of Languages* V.i p.256.

[504] Rousseau, *Essay on the Origin of Languages* XX.iii p.299. Avi Lifschitz, 'How to do Things with Signs: Rousseau's Ancient Performative Idiom', *History of Political Thought* 37 (2016) 46-63 points out that Rousseau thinks that the institution of a republican government also requires the institution of a more performative idiom better able to accommodate the oratory and public gatherings associated with a self-governing society.

empirical objects – or complex ideas of substances as Locke had called them – that allowed them to acquire a generalised, and ultimately even scientific, understanding of the world. Secondly, words allowed humans to construct complex ideas to regulate their conduct – in other words: moral concepts. The ability to compose these two new types of ideas, made possible by language, then allowed for the dramatic expansion of the cognitive powers of nascent man, as well as for his aptitude for sociability. As the invention that made the construction of complex ideas possible, language is therefore responsible for enabling these two perfectible properties of human nature.

Language as the Medium for Prejudice

While language enables humans to construct complex ideas, thereby causing the perfectibility of human nature, Rousseau thinks that this powerful tool also has the potential to lead mankind astray. As we have seen in previous chapters, this theme had been already explored by other philosophers. Many of Rousseau's ideas on the potential detriment of language are thus adaptations from preceding thinkers, providing some different accents rather radical new insights. Rousseau's main worry is that through language, flawed ideas may be composed and disseminated. Thus with the invention of words, man acquired the ability to construct complex ideas without any direct relation to experience, including moral concepts as well as abstract and universal ideas of empirical objects. As we have seen, Rousseau holds a nominalist view that entails that due to their indirect relation to experience, complex ideas are ultimately nothing more than words with a certain definition, and therefore linguistic in nature. But as words not only have a cognitive function, but are also used to communicate ideas, they also spread the ideas they signify throughout society. As a conduit for ideas, language therefore enables the distribution of concepts – both correct and mistaken. Rousseau then refers to the flawed concepts disseminated through language with the term of 'prejudice.'[505]

Although, true to his unsystematic style, Rousseau does not provide a formal definition of the concept of prejudice, we may discern two types of flawed

[505] Rousseau, *Second Discourse* II.liv p.185: 'From the extreme inequality of Conditions and fortunes, from the diversity of passions and talents, from the useless arts, the pernicious arts, the frivolous Sciences, would arise masses of prejudices equally contrary to reason, happiness and virtue.'

ideas that may be composed and spread by the use of words. The first type consists of words that do not have any discernable reference to empirical reality. Thus Rousseau reiterates the point, already made by Hobbes, that words allow man to construct metaphysical ideas that, upon closer inspection, do not have any basis in experience.[506] Consequently, humans are prone to compose and believe in the most fantastical ideas that turn out to be completely absurd and the product of the imagination rather than reason. According to Rousseau, this propensity of metaphysical words to deceive human understanding then explains how modern man has come to believe so many unsound theological and philosophical notions. Possibly inspired by Condillac's *Treatise on Systems*, Rousseau then singles out seventeenth century rationalism, and Descartes in particular, as an example of how the careless use of words can lead to misguided conclusions.[507]

Yet while Rousseau reiterates the concern of previous philosophers for the abuse of metaphysical terms, he is ultimately more worried by the possibility that prejudice corrupts the moral concepts regulating the conduct of individuals in society. In fact, Rousseau thinks that the behaviour of humans living in modern society is almost entirely determined by prejudice:

> Toute nôtre sagesse consiste en préjugés serviles ; tous nos usages ne sont qu'assujetissement, gêne, et contrainte. L'homme civil nait, vit et meurt dans

[506] To be sure, Rousseau himself also occasionally defends metaphysical ideas that are open to this objection. Yet Rousseau is careful not to present these ideas in the context of his main philosophical argument. Thus his main defence of the existence of God and the idea of freedom of the will takes place in the *Confession of the Savoyard Vicar*, a digression in the *Émile* of which it is not entirely clear whether the statements attributed to Rousseau's protagonist are also his own – a question to which we will return below. In any case, as Roger Masters, *The Political Philosophy of Rousseau* (Princeton 1968) p.73-74; Melzer, *The Natural Goodness of Man* 12 maintain, Rousseau's moral and political philosophy should be considered isolated from the metaphysics of the *Confession*, if only because Rousseau's practical philosophy does not appeal to this metaphysics for support. Israel, *Democratic Enlightenment* 102 and David Williams, *Rousseau's Platonic Enlightenment* (University Park 2007) 95 disagree with this position however – each for their own reasons. Thus we will see below that Williams interprets Rousseau as a Platonist with a theory correspondingly anchored in metaphysics. Israel's main goal is to cast Rousseau as a counter-Enlightenment thinker whose reactionary ideas derive from his dualist metaphysics, presenting Rousseau as a main opponent of the monist-materialist *philosophes* he considers as one of the main sources of modern democratic ideals. For an extensive rebuttal of Israel's interpretation of Rousseau see: Helena Rosenblatt, 'Rousseau, the "Traditionalist"', *Journal of the History of Ideas* 77 (2016) 627-635.

[507] Rousseau, *Lettres Morales* II p.1090. Cf. III p.1095.

> l'esclavage : à sa naissance on le coud dans un maillot ; à sa mort on le clouë dans une biére ; tant qu'il garde la figure humaine il est enchaîné pas nos institutions.[508]

In this passage Rousseau refers first of all to social practices subjecting civil man to oppression. But Rousseau also employs the term prejudice to refer to 'nôtre sagesse' – in other words, to our ideas that legitimise the subjection of civilised man. These moral prejudices are likewise sustained and disseminated through the use of words. As moral concepts like 'property,' 'freedom,' or 'justice,' have no clear reference to empirical reality, they are not only linguistic entities, but are also generally acquired from others by means of words. In practice, this means that the moral outlook of a society is determined by moral conventions established with the advent of society and spread by language. Yet as society has evolved into regrettable directions, Rousseau thinks that the moral conventions composed and spread with the help of language have enabled man's sociability only by ensuring his oppression.

Although Rousseau thinks that as an unfortunate consequence of man's perfectibility, prejudice is clearly introduced by means of language, he is somewhat unclear about the precise method by which words spread flawed concepts. The most obvious way by which language enables the dissemination of moral concepts including prejudice is through education. For this reason, the *Émile*, Rousseau's treatise on education, argues for the limitation of the use of language in instruction. Above all, Rousseau prescribes that children should not be exposed to moral or metaphysical terms.[509] This is not only because children are generally unfamiliar with these terms, inducing them to invent their own uninformed significations.[510] Rousseau also thinks that when exposed to moral or metaphysical discourse, children will contract the prejudices of others. For this reason, Rousseau states that for a long time, his pupil would not be allowed to read any books – with the exception of Robinson Crusoe, as its protagonist is a largely solitary and self-sufficient creature.[511]

As an alternative to verbal or written instruction, Rousseau maintains that students should be encouraged to acquire ideas through their own experience. Rousseau therefore maintains that his own pupil would not be taught

[508] Rousseau, *Émile* I p.253.
[509] Rousseau, *Emile* II p.316; II p.421.
[510] Rousseau, *Emile* I p.298; II p.350.
[511] Rousseau, *Emile* III p.455.

scientific or moral knowledge, but would rather be provided the methods to acquire this knowledge himself.[512] Rousseau's model student Émile will then develop his judgment gradually through his own experience and independent from the authority and prejudices of others:

> Forcé d'apprendre de lui-même, il use de sa raison et non de celle d'autrui ; car pour ne rien donner à l'opinion il ne faut rien donner à l'autorité, et la pluspart de nos erreurs nous viennent bien moins de nous que des autres.[513]

According to Rousseau's method of education by experience, Émile will then acquire his own scientific knowledge by investigating natural objects.[514] Yet Rousseau maintains that his pupil should acquire moral concepts by means experience as well, illustrating his point with an example. Thus Rousseau would teach Émile about the concept of property by encouraging him to plant beans on a tract of land which, as it turns out, belongs to someone else. Subsequently the rightful owner ploughs over Émile's crops, which had been planted on a field where rare melons had been sowed. As Émile's beans are destroyed, he experiences the infringement of his own property, meanwhile learning that he needs to respect the property of others. Following these experiences, Émile composes his idea of property himself, rather than acquiring it from others with the help of language.[515]

Furthermore, there is a possibility that Rousseau thought not only that language is responsible for the spread of flawed ideas through its use in education, but also that the linguistic conventions that comprise language itself accommodate prejudice. Like the philosophers discussed in previous chapters, Rousseau considers language as a system of signs with meanings upheld by tacit consent. In order to be understood, speakers are therefore compelled to employ words in line with the standard of common use.[516] In previous chapters, we have seen that some scholars have ascribed to some early-modern thinkers including

---

[512] Rousseau, *Émile* III p.487.

[513] Rousseau, *Émile* III p.486.

[514] Rousseau, *Émile* II p.350: 'Que sert d'inscrire dans leur tête un catalogue de signes qui ne réprésentent rien eux ? En apprenant les choses n'apprendront-ils pas les signes ? Pourquoi leur donner la peine inutile de les apprendre deux fois ? et cependant quels dangereux préjugés ne commence-t-on pas à leur inspirer en leur faisant prendre pour de la science des mots qui n'ont aucun sens pour eux.'

[515] Rousseau, *Émile* II p.331-332.

[516] Rousseau, *Essay on the Origin of Languages* I.i.1 p.248.

Hobbes and Locke the position that in view of the cognitive connection between words and ideas, human beings not only accommodate the signification of terms used in communication to conventional standards, but also model the composition of their ideas – and in particular universal or abstract ideas and moral concepts – on common meanings. Yet we have also determined in previous chapters that this theory of linguistic relativism, according to which human cognition is determined or at least influenced by the structure and content of language, may only be properly ascribed to Condillac.

As Rousseau derives his theory of language and its role in cognition almost entirely from Condillac, we would also expect him to arrive at a position of linguistic relativism. Indeed, Rousseau certainly thinks that the conventions of language have a profound impact on the way in which individuals compose their ideas:

> Mais les langues en changeant les signes modifient aussi les idées qu'ils réprésentent. Les têtes se forment sur les langages, les pensées prennent la teinte des idiomes, la raison seule est commune, l'esprit en chaque langue en partie la cause ou l'effet des caractéres nationaux, et ce qui paroit confirmer cette conjecture est que chez toutes les nations du monde la langue suit les vicissitudes des mœurs et se conserve ou s'altére comme elles.[517]

Like Condillac, Rousseau thus maintains that the conventions of language influence the composition of our ideas. Yet unlike Condillac, who was mostly concerned with the potential of linguistic conventions to interfere with our grasp of metaphysics and empirical reality, Rousseau immediately draws out the consequences of linguistic relativism for morality. Thus Rousseau maintains that the composition of language reflects the moral outlook of its speakers, and consequently that diverging linguistic conventions are an important reason for the moral diversity between nations.[518]

Yet in the end, we should be cautious to ascribe to Rousseau a full-fledged theory of linguistic relativism. Certainly, apart from the passage quoted

[517] Rousseau, *Émile* II p.346.

[518] J. Patrick Dobel, 'The Role of Language in Rousseau's Political Thought', *Polity* 18 (1986) 638-658 maintains that as Rousseau suggests in the *Social Contract* that a self-governing society requires common adherence to conventional mores, Rousseau would therefore favour the wholesale reform of the linguistic conventions that disseminate these mores. While Rousseau's theory of language might suggest this conclusion, Dobel's reading finds little support in the *Social Contract* itself, where Rousseau declines to discuss the role of linguistic conventions at any length.

above, Rousseau's philosophy also incorporates several views that could be presented as circumstantial evidence for such a reading. Thus Rousseau's idea that man's reason has expanded gradually only with the help of the historically developed artefact of language, would also suggest that human cognition is determined to some degree by the way language has evolved. In addition, Rousseau's rejection of moral universalism can be explained as following from the notion that moral concepts are linguistic entities tied to a confined set of conventional meanings.[519] At the same time, the passage quoted above is the only instance in which Rousseau explicitly endorses the position of linguistic relativism. And unlike Condillac, Rousseau does not provide any more extensive account of how conventional standards of signification impact the composition of ideas by individuals – not even in his *Essay on the Origin of Languages*. Thus even if Rousseau at some point endorses a form of linguistic determinism, he appears not to consider the concept significant enough to merit it any further detailed examination.

The Perfectibility of the Passions

As we have seen, Rousseau maintains that the perfectibility of human nature relies on the ability to construct complex ideas, which is in turn made possible by the use of words. Above, we have discussed how this ability significantly enhanced human understanding of the natural world. In addition, we have seen that the ability to construct complex ideas also made it possible to regulate the conduct of members of society through the invention of moral concepts. The

[519] Responding to Diderot's article *Droit Naturel* in the *Encyclopédie*, Rousseau, *Geneva Manuscript* I.ii 161-162 asserts that as man will always favour the common interest of the society of which he is part, it is a delusion to think that human beings will ever recognise or follow a universal morality based on the interest of mankind as a whole. Despite Rousseau's apparent rejection of universal morality in this passage of the *Geneva Manuscript*, scholars have debated at length the question whether Rousseau retained a conception of natural law, or whether he rejected and/or transcended the notion. The first position is taken by Robert Derathé, *Jean-Jacques Rousseau et le science politique de son temps* (Paris 1988) 152-171, while the second is often identified with Leo Strauss, *Natural Right and History* 252-293. In recent decades, a general consensus has emerged that Rousseau indeed retained a certain conception of natural law – even if scholars disagree widely on the question what role this conception plays in Rousseau's moral philosophy. An accessible summary of the recent debate can be found in: Leonard Sorensen, 'Rousseau's fulfilment of the natural public law tradition and his contribution to its demise', *The European Legacy* 10 (2005) 439-454.

scope of human reason is therefore greatly expanded by the acquired ability to construct complex ideas. Yet intellectual reason is not the only perfectible property of human nature. Rousseau maintains that man's passions are augmented in equal proportion to his intellectual faculties. Thus Rousseau maintains that as human understanding developed during the transition from the state of nature to society, man's passions expanded from the desire for things necessary to self-preservation to a craving for countless forms of pleasure. According to Rousseau this development is another unfortunate consequence of man's perfectibility, as it is one of the primary causes of the misery of civilised man.

In order to explain why Rousseau thinks that man's passions expand in conjunction to the development of his intellect, we will first need to understand how Rousseau thinks desires emerge and in what way they determine human actions. In other words, we will need to acquire an understanding of Rousseau's accounts of human motivation and practical deliberation. As with many other elements of his theory of human nature, Rousseau's version of these accounts is clearly inspired by ideas of the thinkers we have discussed in previous chapters. It appears that Hobbes's theory was a particularly important source of inspiration, as Rousseau explicitly mentions and critiques the account of his predecessor.[520] Yet while Rousseau adopts many of the preconceptions of Hobbes's account of practical deliberation, he expands the theory of his predecessor in one crucial respect. Consequently, Rousseau maintains that for the most part, man's passions and fears are not innate, but rather a perfectible element of human nature that is only developed with the evolution of mankind from the state of nature to society.

The fundamental thesis of Rousseau's account of practical deliberation is that all human actions are ultimately motivated by certain desires or aversions – a position we have also encountered in our previous chapters on Hobbes and Locke. Rousseau also maintains that human thoughts are the consequence of our passions, as he states that: 'we seek to know only because we desire to enjoy, and it is not possible to conceive why someone who had neither desires nor fears would take the trouble to reason.'[521] What is more, Rousseau thinks that the

[520] These references have apparently been missed by Douglass, who maintains in *Hobbes and Rousseau* 178-179 that: 'there is little evidence [...] to suggest that his [i.e. Rousseau's] theory of the passions was set out in light of his engagement with Hobbes.'

[521] Rousseau, *Second Discourse* I.xix p.142.

development of human understanding by means of the invention of language was itself occasioned by the desires and fears of nascent man. As we have also seen in previous chapters, the upshot of this position on the question of the origin of human motivation is that reason is no longer conceived as faculty that may independently induce our actions. In line with his conception of reason as the composite of all faculties of human understanding rather than independent faculty, Rousseau thus maintains that it is the passions rather than reason that ultimately motivate all our thoughts and actions.[522]

Yet the fact that all of man's thoughts and actions are ultimately occasioned by certain desires or fears, does not entail that reason does not participate in determining our conduct. In our first chapter, we have seen that Hobbes explains that reason takes part in practical deliberation both by finding the means to our desires, and by foreseeing the consequences of pursuing our passions. Thus while reason cannot initiate actions without being prompted by certain desires or fears, it nonetheless determines which passions are worthwhile pursuing, and by what precise course of action they may be attained. Although Rousseau does not reiterate the details of Hobbes's account, it is clear that he has similar ideas about the role of reason in practical deliberation. Thus Rousseau maintains that the development of reason by means of the invention of language has greatly expanded man's foresight, which in turn has improved his ability to find the means to desires, as well as envision the consequences of his pursuits.[523]

In the first chapter of this thesis we have also seen that Hobbes thinks that for the most part, our desires and aversions do not arise spontaneously, but are rather occasioned by our train of thoughts. Hobbes maintains that apart from a few basic desires generated by physical necessities, our passions emerge only when the mind contemplates an idea of a certain object and considers the prospective pleasurable or painful consequences of pursuing this object. Thus while Hobbes thinks that reason is roused only by the passions, the objects of our desires and fears are nonetheless a consequence of the ideas present in our mind.

---

[522] Rousseau thereby appears to embrace Hobbes's mechanistic account of human motivation, which seems at odds with Rousseau's statement, discussed above, that humans possess free will. Rousseau himself does not address this apparent incongruity.

[523] Rousseau, *Second Discourse* II.v-vi p.162. The way in which the development of reason through language enables the expansion of foresight is not always well understood by Rousseau's interpreters. Zev M. Trachtenberg, *Making Citizens: Rousseau's political theory of culture* (London and New York 1993) 106 even goes so far to claim that 'Rousseau does not explain how men take the crucial step of gaining foresight.'

Again, without recapitulating the details of Hobbes's account, Rousseau agrees with his predecessor:

> The Passions, in turn, owe their origin to our needs, and their progress to our knowledge; for one can only desire or fear things in terms of the ideas one can have of them, or by the simple impulsion of Nature; and Savage man, deprived of every sort of enlightenment, experiences only the Passions of this latter kind; his Desires do not exceed his Physical needs.[524]

Rousseau thus agrees with Hobbes's rather straightforward view that humans can only desire or fear things of which they already have ideas. Yet from this position, Rousseau draws a conclusion that is not made explicit by Hobbes himself – namely that man's desires increase in proportion to his knowledge. The improvement of man's foresight thus not only enhances our ability to find the means to his desires. The expansion of our ideas also multiplies the objects of our desire.[525]

This notion that desires and aversion develop in proportion with our ideas then leads Rousseau to the conclusion man's passions are as perfectible as his ability to reason. Rousseau portrays natural man as an inane and docile creature not because he is wholly without passions, but because he is almost entirely without ideas of objects he may desire or fear. This is the reason why Rousseau maintains in the passage quoted above that the passions of natural man are limited to those necessary for his preservation. The passions of natural man do not extend beyond hunger, thirst, sleep, shelter and the desire to reproduce. Yet with the invention of language, nascent man acquired the ability to compose ever more complex ideas. In addition, by acquiring the use of words, nascent man also further expanded his ideas through communication with those with whom he had started to live in society. The resulting expansion of man's ideas caused an inflation of his passions. Consequently, civilised man started to desire or fear things to which natural man had been oblivious. This is the reason

[524] Rousseau, *Second Discourse* I.xix p.142. Cf. *Émile* II p.304: 'C'est l'imagination qui étend pour nous la mesure des possibles soit en bien soit en mal, et qui par consequent excite et nourrit les desirs pas l'espoir de les satisfaire.'

[525] Scott, 'Rousseau's Unease with Locke's Uneasiness', 305 adds that accordingly, Rousseau diverges from the view of Locke, who had maintained that man is inherently subject to a certain uneasiness that causes passions regardless of man's ideas.

why Rousseau thinks that the passions apparent in modern human beings arose only with the concurrent development of language and society.[526]

The theory that both reason and the passions are not fixed but rather perfectible properties of human nature that develop concurrently, then leads Rousseau to his critique of Hobbes's conception of human nature and the latter's characterisation of life in the state of nature. As we have seen in our first chapter, one of the main reasons why Hobbes's characterised life in the state of nature as a war of all against all, is that in absence of established authority, human beings come into conflict over the many things they desire. Following his theory that man's passions expanded only with the development of reason, Rousseau retorts that Hobbes 'improperly included in Savage man's care for his preservation the need to satisfy a multitude of passions that are the product of Society.'[527] Rousseau rather maintains that the passions and fears that cause conflicts of interest among humans emerged only when they invented language and acquired the ability to construct complex ideas. Accordingly, Rousseau thinks that society, rather than the state of nature, is characterised by conflicts of interest. As natural man desires nothing but things necessary for his physical preservation, Rousseau thinks that without knowing it, humans in the state of nature abide by the natural laws Hobbes had formulated as precepts that ensure man's preservation.[528]

The Passions and the Misery of Civilised Man

Rousseau presents the inflation of desires that occurred during the development of man's faculties as another lamentable consequence of human perfectibility. The reason why Rousseau considers the expansion of the passions as unfortunate is not merely because it has introduced a conflict of interests between human beings, thereby necessitating the institution of the state. Following the inflation of his passions, civilised man has become subject to so many desires and fears that he could not hope to satisfy them all. As we have seen, Rousseau maintains that improved foresight has not only improved man's resourcefulness, but also greatly expanded his passions. For this reason, Rousseau thinks that the

[526] Rousseau, *Émile* IV p.490-491.
[527] Rousseau, *Second Discourse* I.xxv p.151.
[528] Rousseau, *Second Discourse* I.38 p.154.

expansion of foresight has not contributed to the happiness, but rather to the misery of civilised man:

> La prévoyance! La prévoyance qui nous porte sans cesse au-delà de nous et souvent nous place où nous n'arriverions point ; voila la véritable source de toutes nos misères.[529]

In his typical elliptical style, Rousseau explains why increased foresight is often a curse rather than a blessing. Following the development of human understanding, civilised man becomes aware of opportunities for pleasure and prospects of pain so numerous he will never be able to appease all passions resulting from this improved foresight. Consequently Rousseau thinks that in the face of so many unfulfilled passions, civilised man will remain perpetually miserable.[530]

By arriving at this last conclusion, Rousseau implicitly rejects the conception of felicity retained by his predecessors including Hobbes. In our first chapter, we have seen that Hobbes, who thought that unrelenting passions are an inescapable part of the human condition, maintained that felicity consisted in the continuous satisfaction of successive desires. According to Hobbes, happiness is therefore a transient condition that depends on our ability to find the means to our passions as well as foresee the consequences of our actions. For this reason, Hobbes seems to suggest that felicity would be most effectively pursued through the development of reason leading to extended foresight. But Rousseau recognises that extended foresight not only allows us to satisfy our passions, but also makes us discover new objects of desire. Thus even if civilised man is adept at finding ways to satisfy his desires, his expanded foresight will generate ever more passions. He therefore finds it unlikely that expanding our reason will make us happy.[531] Afflicted by unrelenting appetite, Rousseau thinks that the happiness of civilised man is not merely transient, but entirely non-existent.

As an alternative to Hobbes's conception of felicity, Rousseau formulates his own account of happiness. To begin with, Rousseau agrees with Hobbes that

---

[529] Rousseau, *Émile* II p.307.

[530] Rousseau, *Émile* II p.311-312. This is also one of the reasons why Rousseau, in his *First Discourse*, defends the thesis that the advance of the arts and sciences had corrupted rather than purified man's morals. Thus Rousseau maintains that human progress has made available a range of temptations that have occasioned both man's misery and vice.

[531] Rousseau, *Émile* II p.304.

man will be miserable when unable to satisfy his passions and achieves happiness by fulfilling his desires. Yet Rousseau maintains that felicity is achieved not merely by simply extending foresight, but rather by ensuring that our passions do not exceed our power to satisfy them:

> Tout sentiment de peine est inséparable du desir de s'en délivrer ; toute idée de plaisir est inséparable du desir d'en joüir ; tout desir suppose privation, et toutes les privations qu'on sent sont pénibles ; c'est donc dans la disproportion de nos desirs et de nos facultés que consiste nôtre misére. Un être sensible dont les facultés egaleroient les desirs seroit un être absolument heureux.[532]

For Rousseau, happiness thus consists in an equilibrium in which foresight is sufficiently developed to ensure the satisfaction of our present desires, but not amplified to the point of invoking passions beyond our grasp.[533] Rousseau therefore states that happiness is achieved by 'diminuer l'excés des desirs sur les facultés, et à mettre en égalité parfaite la puissance et la volonté.'[534]

Rousseau presents natural man as the archetype of a being endowed with powers equal to his desires. As we have seen, Rousseau maintains that natural man has very few ideas, and consequently desires nothing but the things absolutely necessary to his preservation. Furthermore, we have also seen above that Rousseau characterises the state of nature as a state of plenty in which natural man has little difficulty in obtaining the necessities of life. As the objects of natural man's desire are all within his grasp, his passions are equal to his ability to satisfy them. Contrary to Hobbes, Rousseau therefore thinks that the state of nature is as close to happiness as man can get.[535] Undisturbed by the various passions that will plague civilised man, natural man enjoys a state of

---

[532] Rousseau, *Émile* II p.303-304. Cf. III p.426.

[533] Christopher Brooke, *Philosophic Pride: Stoicism and Political Thought from Lipsius to Rousseau* (Princeton 2012) 189-192 maintains that Rousseau's conception of felicity is inspired by the Stoic notion of happiness as the limitation of desires to things that are within our power. Yet at the same time Brooke is aware that Rousseau's conception of felicity is derived from a clearly Epicurean account of the relationship between reason and the passions.

[534] Rousseau, *Émile* II p.304.

[535] Rousseau, *Émile* II p.308. Although in the *Social Contract* I.viii p. 56, Rousseau also maintains that the transformation of natural man into civil man is a happy event, as man is no longer a slave to his appetite man and gains freedom through obedience to a law one has himself prescribed. Yet presumably, this happiness is achieved only when living under the constitution described in the *Social Contract* in which the only constraints to citizens are self-imposed, which is not the case with civilised man as found in modern society.

tranquillity in which he experiences nothing more than a sweet sentiment of existence.[536] For Rousseau, the figure of natural man thereby functions as an ideal that we should try to approximate in order to achieve our own happiness.[537]

Conversely, Rousseau thinks that when leaving the state of nature, man will generally be unable to retain this state of felicity. We have already seen that the invention of language and concurrent development of man's faculties expands his passions – generally beyond his powers. But Rousseau thinks that man's pursuit of felicity is further impaired by living in society, as civilised man has become almost entirely dependent on society to obtain the objects he has come to desire. Unlike natural man, who had been self-sufficient in the satisfaction of his limited desires, civilised man requires the cooperation of his fellows in order to fulfil all his acquired passions. Often this cooperation will be difficult to obtain, further frustrating civilised man in the pursuit of his desires.[538] In order to overcome this difficulty, civilised man has developed a desire for power, as it would seem to enable him to compel his fellows into cooperating in the satisfaction of his desires. Yet Rousseau recognises that the acquisition of power – in any form whatsoever – will never truly enable civilised man to escape the dependency on society. Not only is power never truly absolute, Rousseau also maintains that by exercising power, man will in fact become more dependent on those he subjugates.[539]

In view of these properties of society, the question arises how Rousseau thinks felicity conceived as equilibrium between passions and foresight may in fact be achieved by civilised man. To begin with, Rousseau thinks it would be futile to attempt to simply suppress or ignore our passions, as he considers desires and fears as an unavoidable part of human nature.[540] Nonetheless, Rousseau thinks there is at least one way in which passions and powers could be

[536] It is beyond this chapter to conduct a closer discussion of Rousseau's concept of the sentiment of existence. In short, this sentiment is experienced when one has extinguished all passions and has no concern for past or future – in other words when one has achieved a state of perfect happiness. For an extended discussion of this sentiment and the various ways in which scholars interpret its role in Rousseau's thought, see Eve Grace, 'The Restlessness of "Being": Rousseau's Protean Sentiment of Existence', *History of European Ideas* 27 (2001) 133-151.

[537] Rousseau, *Émile* II p.303-304.

[538] Rousseau, *Émile* II p.311-312.

[539] Rousseau, *Émile* II p.308. Melzer, *The Natural Goodness of Man* p.62-63 notes that Rousseau therefore does not oppose oppression because he thinks it violates human dignity, as a Kantian reading would suggest, but rather because it enslaves both slave and master. Cf. p.73-74.

[540] Rousseau, *Émile* IV p.490-491.

balanced in civilised man. Rousseau presents this method in his *Émile*, where he describes a method of education according to which the pupil is to be fashioned according to the ideal of natural man. Rousseau aims to approximate this ideal first of all by shielding his pupil from the rest of society as much as possible, thereby limiting Émile's ideas of objects he may otherwise had come to desire. At the same time, Rousseau aims to fashion Émile into a largely self-sufficient being, capable of acquiring for himself the objects of his limited passions. As a result, Émile will experience few artificial passions, while he will generally be able to satisfy his desires without depending on society.[541]

Ultimately it is not clear to what extent Rousseau himself thinks the education of Émile could be a viable method of bringing all humans living in society closer to the ideal of self-sufficient natural man.[542] Consequently, it is not certain whether Rousseau thought that the ideal of achieving equilibrium between passions and abilities could ever be achieved by civilised man. Yet Rousseau's conception of felicity as the equilibrium between the passions and the power to satisfy them nonetheless has clear implications for his conception of morality and his views on the purpose of society. As Grace Roosevelt has noted, Rousseau's notion of felicity excludes both a consequentialist conception of morality and a utilitarian view on the purpose of society. Not only is it opposed to Hobbes's view of morality as a prudential guideline to the most effective and durable satisfaction of desires. Rousseau's notion of felicity also provides grounds to question Helvétius's conclusion that the purpose of society is to provide the greatest happiness for the greatest number, conceived simply as the maximisation of pleasure among its citizens. As we have seen, Rousseau maintains that happiness is not simply dependent on satisfying our passions, but also on limiting our desires to objects within our grasp. Thus while utilitarian and consequentialist conceptions of morality may direct man towards pleasure, they do not seek to prevent the inflation of the passions, thereby ultimately failing to secure man's felicity.[543]

---

[541] Rousseau, *Émile* II p.311.

[542] Rousseau, *Emile* I p.248 maintains that one must choose between fashioning man into a citizen, or into a self-sufficient being.

[543] Grace Roosevelt, 'An Alternative to Economic Man: The Limitation of Desire in Rousseau's *Émile*', in: Mark Blackell, John Duncan, Simon Kow (eds.), *Rousseau and Desire* (Toronto 2009) 46-61, 53

From Amour de Soi to Amour Propre

In the preceding sections, we have discussed Rousseau theory of the perfectibility of human nature by exploring the consequences of the expansion of man's intellectual faculties and the extension of his foresight. Yet Rousseau thinks that the emergence of society and concurrent development of reason also transforms the way in which individuals relate themselves to others. Following the development of sociability and social hierarchy, civilised man acquires a certain position or status relative to other members of society. In addition, as Rousseau maintains that reason or judgment enables man to make comparisons, the improvement of this ability also creates the capacity to compare his own standing to that of others. Conscious of his own station and motivated by self-love, civilised man will then be inclined to improve his status in comparison to that of his fellows by seeking honour and recognition. Yet Rousseau thinks that this pursuit will necessarily be frustrated, as civilised man will never receive recognition from society sufficient to appease his desire for honour and glory. Consequently, Rousseau maintains that the ability to appreciate moral relations is another unfortunate result of man's perfection caused by the development of reason, as it contributes to the misery more than to the happiness of civilised man.

Rousseau himself describes the transformation described above as the development of *amour de soi*, or self-love, into *amour propre* – a term usually not translated, as there is no English equivalent that entirely captures its meaning. According to Rousseau, every animal, including natural man, is moved by a sentiment of self-love. This *amour de soi* is the origin of their desire for self-preservation.[544] Yet Rousseau maintains that in society, the benign sentiment of *amour de soi* is transformed into the harmful feeling of *amour propre*:

> Self-love is a natural sentiment which inclines every animal to attend to its self-preservation and which guided in man by reason and modified by pity, produces humanity and virtue. Amour propre is only a relative sentiment, factitious, and born in society, which inclines every individual to set greater store by himself than by anyone else, inspires men with all the evils they do one another, and is the genuine source of honour.[545]

[544] Rousseau, *Émile* IV p.491.
[545] Rousseau, *Second Discourse* note XV.i p.218.

By differentiating *amour propre* from *amour de soi*, Rousseau again describes certain qualities of civilised man as perfectible that had previously been considered by many as innate properties of human nature. Thus in our first chapter, we have seen that Hobbes maintained that in their pursuit of power, human beings are naturally disposed to desire glory and honour. Hobbes then identifies this desire for social advancement as a primary cause of the conflicts of the state of nature. By contrast, Rousseau maintains that man's *amour propre*, and the resulting desire for social recognition, emerges only with the advent of society and development of reason.

The most obvious reason why Rousseau thinks that only civilised man is subject to *amour propre* is that he describes natural man as an entirely solitary being. Consequently, natural man need not – and in fact cannot – concern himself with his social status in the eyes of his fellows. Secondly, Rousseau thinks that *amour propre* requires the development of reason. Rousseau does not explain in detail how *amour propre* is developed through reason. Yet Rousseau's account of reason does provide us with an important indication why the ability to reason engenders *amour propre*. Thus above we have seen that Rousseau describes reason as the ability to determine the relations between ideas by means of comparison. Yet Rousseau thinks that man requires this same ability to ascertain his relations to other members of society and compare his own social status to that of his fellows. As the faculty that makes this comparison possible, Rousseau refers to reason as the guide of *amour propre*.[546]

According to Rousseau the development of *amour propre* out of *amour de soi* is another unfortunate consequence of man's perfectibility. The first reason why the development of *amour propre* contributes to the misery of civilised man is already indicated in the passage quoted above, where Rousseau claims that civilised man is incited to compete for honour. Like Hobbes, Rousseau thinks that this competition for social status will be a primary cause for conflicts between individuals. But unlike Hobbes, who thought that these engagements would be especially prevalent in the state of nature, Rousseau thinks that the competition for social status and resulting conflicts would only develop with the advent of society, and therefore only affect civilised man. Secondly, Rousseau maintains that *amour propre* makes civilised man dependent upon the judgment of his

[546] Rousseau, *Émile* II p.322. Cf. Rousseau, *Second Discourse* I.xxxvii p.153.

fellows.[547] According to Rousseau, this desire for esteem also contributes to the misery of civilised man as the desire for social recognition can never be entirely satisfied:

> L'amour de soi, qui ne regarde qu'a nous, est content quand nos vrais besoins sont satisfaits ; mais l'amour-propre, qui se compare, n'est jamais content et ne sauroit l'être, parce que ce sentiment, en nous préférant aux autres, éxige aussi que les autres nous préférent à eux, ce qui est impossible. Voila comment les passions douces et affectueuses naissant de l'amour de soi, et comment les passions haineuses et irascibles naissant de l'amour-propre.[548]

Rousseau explains that in its essence, *amour propre* is a sentiment that causes us to value ourselves over others. In order to appease this sentiment, civilised man seeks signs of recognition that confirm his feeling of superiority. Yet as Rousseau notes, for this reason *amour propre* can only be pacified if others signify they value us above themselves. But this is unlikely to happen, as members of society are all equally subject to *amour propre*. Consequently, Rousseau thinks that signs of social recognition are never sufficient to entirely fulfil our desire for honour.

In Rousseau's view, the development of *amour propre* has thereby contributed to the misery of civilised man much in the same way as the inflation of desires has done. Accordingly, Rousseau identifies the development of *amour propre*, together with the inflation of the passions, as the two primary causes of the misery of civilised man.[549] On the one hand, *amour propre* has transformed us from self-sufficient creatures into beings dependent upon the social recognition of our fellows. On the other hand, like in the case of other passions, the desire for honour and social recognition will never be entirely satisfied, thereby impeding the happiness of civilised man. Rousseau's solutions to the problems caused by the development of *amour propre* resemble those presented to counter the consequences of the inflation of the passions.[550] Thus in the *Émile*,

---

[547] Rousseau, *Second Discourse* II.lvii p.187: 'sociable man, always outside himself, is capable of living only in the opinion of others and, so to speak, derives the sentiment of his own existence solely from their judgment.'

[548] Rousseau, *Émile* IV p.493.

[549] Rousseau, *Émile* IV p.493 : 'Ainsi ce qui rend l'homme essentiellement bon est d'avoir peu de besoins et de peu se comparer aux autres ; ce qui le rend essentiellement méchant est d'avoir beaucoup de besoins et de tenir beaucoup à l'opinion.'

[550] For this reason, and in the interest of brevity, we have kept our discussion or Rousseau's solutions to the problems of *amour propre* rather short. We will briefly return to this issue in a

Rousseau explains how the development of *amour propre* may be limited by shielding pupils from wider society until a certain age and making them unaccustomed to the esteem of others.[551]

Pity and Conscience

In previous chapters, we have seen that Hobbes, Locke, Condillac and Helvétius all presented morality as a prudential rule, found out by reason, indicating the most effective and durable means to satisfy our self-interested desires.[552] Yet in our preceding discussion of Rousseau's conception of felicity, we have seen that Rousseau does not think that by itself, the satisfaction of our passions will lead us to happiness. Accordingly, Rousseau rejects the notion of his predecessors that morality should be the product of prudential reasoning. In fact, Rousseau maintains that by itself, no form of reasoning may be able to establish genuine moral guidelines:

> Par la raison seule, indépendamment de la conscience, on ne peut établir aucune loi naturelle; et que tout le droit de la Nature n'est qu'une chimere, s'il n'est fondé sur un besoin naturel au coeur humain.[553]

Rather, Rousseau thinks that reason may only be able to perceive morality when guided by a sentiment of pity that is somehow innate to human nature. This sentiment therefore performs a crucial function in Rousseau's moral epistemology. According to Rousseau, the sentiment of pity is aroused when man perceives another suffering being. Due to this sentiment innate to human nature, man has the tendency to sympathise with the plight of others. Rousseau

---

subsequent section of the *Social Contract* as a potential solution to the problems caused by the perfection of human nature. For a more extensive discussion of the concept of *amour propre* and Rousseau's remedies to the problems caused by its development, see: N.J.H. Dent and Timothy O'Hagan, 'Rousseau on "Amour-Propre"', *Proceedings of the Aristotelian Society* 99 (1999) 91-107

[551] Rousseau, *Émile* IV p.504.

[552] Although we have seen in a previous chapter that Locke's moral philosophy may be interpreted differently.

[553] Rousseau, *Émile* IV p.523; Cf. p.599, where Rousseau questions whether self-interested reason would be able to convince us to behave morally, even if this would be against our own best interests. Rousseau thereby appears agree with Hobbes's fool, who had posed the same question. *Second Discourse* preface 6 p.126 adds that society could not have been established if morality is only accessible through philosophical reasoning, simply because primitive man did not have the intellectual capacity for such reasoning, as even most modern humans lack the ability for philosophical reflection.

therefore maintains that pity is generally the cause of various forms of virtuous behaviour[554]

Rousseau thereby appears to propose a version of moral sense theory, according to which morality is not the product of reason alone, but rather of a non-cognitive sentiment innate to human nature. Indeed, Rousseau was not the first early-modern thinker to hold this position. In fact, his theory of pity can be related to a wider early-modern tradition that considered morality as the product of innate sentiments. This school of thought includes British thinkers such as Shaftesbury, Hutcheson, Hume, and Smith.[555] But as Helena Rosenblatt has pointed out, it is likely that Rousseau had also encountered versions of moral sense theory in the works of Swiss natural lawyers like Burlamaqui and Barbeyrac.[556] Yet we will see in the following that despite its affinity with previous moral sense theories, Rousseau's sentiment of pity is not simply an autonomous moral sense that would indicate man the morality of his actions – even if it has sometimes been interpreted as such by modern scholars. In fact, we will discover that Rousseau takes care to accommodate his theory of pity to his broader theory of human nature founded on the empiricist principles he derived from the authors we have discussed in previous chapters. Thus Rousseau thinks that in civilised man the sentiment of pity operates only in conjunction with of reason. But as pity relies on the perfectible faculty of reason, we will see that man's moral sense or conscience turns out to be a perfectible faculty as well.

Unfortunately Rousseau's remarks on the sentiment of pity and the way in which it governs our behaviour are somewhat disparate and therefore open to

---

[554] Rousseau, *Second Discourse* I.xxxvii p.153 : 'Mandeville clearly sensed that, for all their morality, men would never have been anything but monsters if Nature had not given them pity in support of reason: but he did not see that from this single attribute flow all the social virtues he wants to deny men. Indeed, what are generosity, Clemency, Humanity, if not Pity applied to the weak, the guilty, or the species in general? Even Benevolence and friendship, properly understood, are the products of a steady pity focused on a particular object; for what else is it to wish that someone not suffer, than to wish that he be happy?'

[555] See Michael Frazer, *The Enlightenment of Sympathy: Justice and the Moral Sentiments in the Eighteenth Century and Today* (Oxford 2010) for an extended discussion of this British tradition, focusing primarily on Hume and Smith.

[556] Helena Rosenblatt, *Rousseau and Geneva: From the First Discourse to the Social Contract 1749-1762* (Cambridge 1997) 93-98. For a discussion of Barbeyrac's moral sense theory, see: Tim Hochstrasser, 'Conscience and Reason: The Natural Law Theory of Jean Barbeyrac', *The Historical Journal* 36 (1993) 289-308.

interpretation.[557] In some passages, Rousseau appears to maintain that the existence of the sentiment of pity entails that human beings are naturally endowed with a conscience that would indicate to them the morality of their actions:

> Il est donc au fond des ames un principe inné de justice et de vertu, sur lequel, malgré nos propres maximes, nous jugeons nos actions et celles d'autrui comme bonnes ou mauvaises, et c'est à ce principe que je donne le nom de conscience.[558]

Some scholars take passages as the one quoted above at face value, and attribute to Rousseau the view that conscience is innate to human nature. The most notable version of this reading is proposed by David Williams, who ascribes to Rousseau a Platonic theory of morality. According to Williams, Rousseau's conscience should be interpreted as a sentiment giving us knowledge of justice in a similar fashion as reason provides a transcendent idea of justice in Plato's theory.[559] Williams even goes so far as to claim that accordingly Rousseau is 'the most consistent and greatest Platonist of the modern era.'[560] This last statement should probably be read as exaggerated for rhetorical purposes. Yet even when Williams' interpretation is read charitably, it remains problematic for at least three reasons.

First of all, the differences between Plato's and Rousseau's moral philosophies are too substantial to warrant close identification. Thus while both certainly appear to ascribe to man some innate sense of morality, Plato maintains that it is accessible only by cultivating an innate faculty of reason, while Rousseau maintains that man's sense of morality originates in a sentiment of pity that, as we shall see below, he thinks is in fact both stifled and perfected by the development of reason. Secondly, Williams' reading is contradicted by Rousseau's

---

[557] Hanley, 'Rousseau's Virtue Epistemology' p.255 also notes the disparity between Rousseau's statements on conscience. Shklar, *Men and Citizens* p.38 maintains that Rousseau in fact changed his mind and included an innate moral sense in his later writings including the *Émile*. Yet as we will demonstrate below, the *Émile* itself contains passages in conflict with this interpretation.

[558] Rousseau, *Émile* IV p.598.

[559] Williams, *Rousseau's Platonic Enlightenment* 75. Disputed by Melzer, *The Natural Goodness of Man* 20-21.

[560] Williams, *Rousseau's Platonic Enlightenment* 95. For a more fruitful comparison between Rousseau's conception of practical reason and those of Plato and Aristotle, see Marshall, 'Epistemology and Political Perception', 97-99, who indeed concludes that Rousseau's conception diverges markedly from that of his ancient predecessors.

unequivocal contention that morality or natural law is in no way innate to human nature, as well by the claim that by itself, the sentiment of pity does not provide us with any concrete knowledge of morality.[561] Thirdly, Williams' reading also relies heavily on passages of which it is not certain that they precisely represent Rousseau's view. Thus the passage quoted above is taken from the *Confession of the Savoyard Vicar*, a section in the *Émile* where Rousseau not explicitly expounding his own theory, but rather presents the views of a fictional character identified only as a vicar from Savoy. Rousseau is equivocal about the question whether the vicar's views are actually his own, giving rise to debate among modern scholars on the issue.[562] On the one hand, Rousseau maintains at some point that the *Confession* provides an outline of his personal beliefs.[563] Yet on the other hand, Rousseau also states in a passage directly following the *Confession* that he does not agree with everything the vicar had to say, suggesting that he disagrees with some of the details of the philosophy elaborated in the *Confession*.[564]

Yet the most significant argument against William's reading is the fact that elsewhere in his writings, Rousseau provides a different and more detailed account of the way in which conscience develops out of an innate sentiment of pity. Rather than as an autonomous moral sense, Rousseau here describes how pity gives rise to conscience in conjunction with man's intellect:

> The social affections develop in us only with our knowledge. Pity, although natural to man's heart, would remain eternally inactive without imagination to set it in motion. How do we let ourselves be moved to pity? By transporting ourselves outside ourselves; by identifying with the suffering being. [...] think how much acquired knowledge this transport presupposes! How could I imagine evils of which I have no idea?[565]

[561] Rousseau, *Geneva Manuscript* 160-16; *Émile* I p.288.

[562] Williams, *Rousseau's Platonic Enlightenment* 62-64 maintains that the Vicar indeed is a mouthpiece for Rousseau himself, while Melzer, *The Natural Goodness of Man* 30n; 147 contests this view.

[563] Rousseau, *Reveries* in: *Œuvres* I p.1018; *Lettres écrites de la montagne* in: *Oeuvres* III p.694.

[564] Rousseau, *Émile* IV p.606.

[565] Rousseau, *Essay on the Origin of Languages* I.ix.2 p.267-268. See for comparable passages: *Émile* IV p.505-506; *Lettre a Beaumont* p.936.

Rousseau here maintains that rather than a fully formed moral sense or conscience, it is only the sentiment of pity that is innate to human nature.[566] Yet pity is only activated when man identifies with the suffering being he has encountered by putting himself in its place. Accordingly, the process of identification requires imagination, which in turn is developed only with the growth of knowledge – or the perfection of reason. Rousseau thus maintains that while the repugnance to seeing another being suffer is innate to the human heart, it is only truly activated when man acquires enough knowledge to imagine the suffering of those he encounters.[567]

Yet this conception of conscience as requiring both the sentiment of pity and the perfection of man's intellectual faculties also raises an important question. At first glance, it would suggest that the development of reason would also cause the perfection of conscience. But then why is there no evidence that civilised man indeed has a conscience more developed than that of natural man? To begin with, Rousseau thinks that man's development from the state of nature to society, and the intellectual awakening that accompanied it, have also interfered with man's conscience in several ways. Thus Rousseau maintains that the conscience of civilised man is often suppressed by prejudices, which as we have seen are disseminated throughout society with the help of language.[568] More importantly, Rousseau thinks that man's conscience is also stifled by the development of *amour propre*, which makes civilised man care for his self-interest more than the well-being of others. We have seen that Rousseau maintains that following the emergence of society, civilised man becomes concerned with his place in the social hierarchy and starts to desire the esteem of his fellows. Rousseau explains that the social hierarchy resulting from the emergence of society and the development of *amour propre* also reduces the scope of man's conscience. Thus Rousseau thinks that man will have difficulty identifying with individuals situated at different rungs of the social ladder. Thus civilised man will generally have contempt for his inferiors, while envying those with greater achievements. In addition, he will generally be unfamiliar with the troubles that afflict those in other social classes. The development of *amour*

[566] Rousseau, *Second Discourse* I.xxxv p.152 maintains in fact that even animals appear to possess this sentiment. But as animals lack the faculty of reason, they do not develop conscience or any sense of morality.

[567] Hanley, 'Rousseau's Virtue Epistemology', 255-256 agrees with this interpretation.

[568] Rousseau, *Émile* I p.245.

propre and the rise of social distinctions have therefore almost entirely silenced the sentiment of pity in civilised man.[569]

In addition, Rousseau thinks that the impact of the sentiment of pity is altered as the development of reason expands the scope of man's conscience. Thus Rousseau explains that natural man may be moved by pity, but only by the suffering of individuals identical to himself.[570] The development of reason and the growth of imagination causes civilised man to identify and commiserate with many more individuals, even those that show no immediate similarity to him. Furthermore, intellectual reason even allows man to commiserate with the evils in society as a whole. Yet Rousseau maintains that the sentiment of pity is only faintly aroused when instructed by reason. Thus Rousseau provides an example of a philosopher who is kept up at night by the problems of society, but who remains unmoved when someone is murdered beneath his window. According to Rousseau, the development of reason thus broadens man's conscience, but may at the same time desensitise him to acts that would have been extremely repugnant to natural man. The conscience of civilised man is thus greatly expanded in scope, but at the same time it has lost much of its original force and vivacity. Rousseau therefore maintains that pity is 'a sentiment that is obscure and lively in Savage man, developed but weak in Civilised man.'[571]

Our extended account of Rousseau's theory of conscience thus indicates that it is not an autonomous moral sense, but rather a perfectible faculty that arises out of the combined effort of pity and the intellect. For this reason, it also follows that with man's cognitive development and the institution of society, the operation of his conscience must have been modified as well. Indeed, Rousseau maintains that in the state of nature pity is still forceful enough to make natural man behave morally. Accordingly, Rousseau states that in the state of nature 'pity takes the place of Laws, morals, and virtue, with the advantage that no one is tempted to disobey its gentle voice.'[572] Yet while the development of reason expanded man's conscience in scope, the transformation from the state of nature to society and emergence of *amour propre* have also impaired the force of the sentiment of pity to the extent that it is easily ignored by civilised man. This weakening of pity coinciding with man's development is then one of the primary

[569] Rousseau, *Émile* IV p.503-505; 508-509; p.514; Shklar, *Men and Citizens* p.47.
[570] Rousseau, *Second Discourse* I.xxxvii p.153.
[571] Rousseau, *Second Discourse* I.xxxvii p.153.
[572] Rousseau, *Second Discourse* I.xxxviii p.154.

reasons why Rousseau thinks man is naturally good and has only become evil with the emergence of society.

Finally, our account of Rousseau's conception of conscience as a compounded and perfectible faculty rather than an innate moral sense is also accommodated to his general theory of the perfectibility of human nature. Moreover, we have tried to show in the preceding that this theory was inspired to considerable degree by previous empiricist theories of human nature. Our reading presenting Rousseau's conception of conscience not as an innate moral sense but as a faculty arising out of the combined operation of pity and imagination complies with one of the general predispositions of empiricism – the tendency to ascribe human nature as few innate properties as possible. Although Rousseau undoubtedly claims that the sentiment of pity is innate to human nature, the activity of this sentiment relies on the imagination – a cognitive faculty considered part of human nature by all empiricists. By describing pity merely as a repugnance to seeing other beings suffer, which is only developed into conscience with the help of the imagination, Rousseau therefore ascribes to man a form of moral sensitivity without departing too much from either the empiricist prohibition against assuming the existence of innate ideas, or from the inclination to ascribe to human nature as few natural faculties as possible.

Human Nature and Morality

At the end of our exploration of Rousseau's theory of human nature, we are now in the position to outline how Rousseau's conception of man provides the basis for his moral philosophy. In previous chapters we have seen that following Hobbes, Rousseau's predecessors employed the thought experiment of the state of nature to demonstrate how in view of their theoretical account of human nature, individuals interact with each other without established morality or centralised authority. They subsequently concluded to a varying extent that life in the state of nature would be characterised by conflict and hardship. Again with their conception of human nature in mind, Hobbes and his followers subsequently designed prudential rules that would remedy this situation, thereby ensuring the collective safety and prosperity of mankind. In our discussion of Rousseau's account of human development at the start of this chapter, we have already seen that Rousseau considers these efforts of his predecessors as misguided. Rousseau has many objections, but his most important reason to reject their accounts of the state of nature is that he thinks human nature is

perfectible and has transformed fundamentally during the development from nature to society.

Yet this does not entail that Rousseau rejects the thought experiment of the state of nature as a useful approach in the search for a theory of morality. Instead, Rousseau reformulates the thought experiment to account for his conception of human nature as perfectible rather than static and universal. As we have seen, Rousseau considers the state of nature not simply as a condition without society, law, and centralised authority. Rather, he describes it as man's original station, not only predating the emergence of society, but also the concurrent development of language and man's perfectible faculties. We have seen above that Rousseau thinks that in the state of nature mankind must have had a simple yet felicitous existence. Tacitly assuming that Nature or God would not have created man unsuited to his natural state, Rousseau seems to suggest that life in the state of nature is also suited to man's nature.[573] For Rousseau the thought experiment of the state of nature therefore no longer serves to demonstrate the rules necessary to contain man's destructive tendencies for the collective welfare – simply because these tendencies are not yet present in natural man. Instead, Rousseau's state of nature displays a type of existence suited to man's nature. For this reason, Rousseau's moral philosophy presents the properties and way of life of natural man as an ideal that human beings in all states of development should strive to emulate.[574]

Yet this notion of natural man as an ideal raises another question: in what way can man in his natural state provide an ideal for the morality of civilised man, whose nature has been irreversibly altered? We have already seen that in contrast to what some of his contemporary critics may have suggested,

[573] Rousseau, *Émile* II p.322: 'Posons pour maxime incontestable que les premiers mouvemens de la nature sont toujours droits: il n'y a point de perversité originelle dans le cœur humain. Il ne s'y trouve pas un seul vice dont on ne puisse dire comment et par où il y est entré.' Cf *Lettre à Beaumont* p.935-936.

[574] Rousseau, *Second Discourse* preface 11 p.128: 'This same study of original man, of his true needs, and of the fundamental principles of his duties is also the only effective means available to dispel the host of difficulties that arise regarding the origin of moral inequality, the true foundations of the Body politic, the reciprocal rights of its members, and a thousand similar questions, as important as they are badly elucidated.' Delaney, *Rousseau and the Ethics of Virtue* 137-138 provides a more exhaustive list of all the properties of natural man that should be emulated by civilised man.

Rousseau does not simply intend man to return to his primitive condition.[575] Rather, Rousseau identifies the properties of natural man that made life in the state of nature peaceful and felicitous, and subsequently devises a moral philosophy that attempts to rediscover, revive, or reconstruct these properties in civilised man. Accordingly, Rousseau presents his moral philosophy as recapturing two properties of natural man suppressed following the development of reason:

> Meditating on the first and simplest operations of the human Soul, I believe I perceive in it two principles prior to reason, of which one interests us intensely in our well-being and our self-preservation, and the other inspires in us a natural repugnance to seeing any sentient Being, and especially any being like ourselves, perish or suffer. It is from the cooperation and from the combination our mind is capable of making between these two Principles, without it being necessary to introduce into it that of sociability, that all the rules of natural right seem to me to flow; rules which reason is subsequently forced to re-establish on other foundations, when by its successive developments it has succeeded in stifling Nature.[576]

Rousseau maintains that before the development of reason, natural man is motivated solely by self-love and has not yet acquired *amour propre*. Accordingly, Rousseau also thinks natural man still experiences the sentiment of pity when he sees other humans suffer. We have already seen that Rousseau maintains that consequently, natural man unconsciously behaves in accordance with natural right. In the passage quoted above, Rousseau suggests that it is the task of reason or philosophy to reformulate these natural properties into a theory of morality for the benefit of civilised man.

But what does this entail in practice? Throughout our preceding discussion, we have already encountered some measures by which Rousseau intends to reform civilised man and negate some of the negative effects of living in society. These measures can be categorised into two distinct projects, each providing a different solution to the problems that have arisen following the perfection of human nature and development of society. The first of these projects, presented in the *Émile*, provides various means by which infants may be

[575] In response to the *Second Discourse*, Voltaire wrote Rousseau a scathing letter in which he proclaimed that Rousseau had written a book against the human race and that it had made him want to return to walking on all fours. See: Voltaire to Rousseau, 30 aug 1755 in: *Correspondance Complète de Jean-Jacques Rousseau* ed. R.A. Leigh (Oxford 1965-1998) III pp.156-157.

[576] Rousseau, *Second Discourse* preface ix p.127.

educated to become moral, self-sufficient and felicitous human beings in modern society. As man's perfectibility has enabled the transformation of human nature with the development of language and advent of society, this same perfectibility also allows Rousseau to devise a method of education that avoids or counteracts many of the defective habits and prejudices usually contracted by children growing up in modern society. Throughout our discussion, we have encountered several examples of measures by which Rousseau intends to shield his pupil from prejudice, stimulate the development of conscience, as well as limit the development of *amour propre*. As a result, *Émile* will become a human being moved by pity, but motivated by neither the desire for the esteem of others, nor by passions inflated beyond his ability to satisfy. Just like natural man, *Émile* will then become self-sufficient, conscientious and content with the necessities required for survival.[577] For this reason, Rousseau claims that his goal is to fashion *Émile* into 'un sauvage fait pour habiter les villes.'[578]

The Political Solution of the Social Contract

Apart from his program for the reform of the individual by means of education outlined in the *Émile*, Rousseau also provides a solution to the problems of civilised life that may be applied to society as a whole. The most comprehensive statement of this solution may be found in Rousseau's *On the Social Contract*. There is a tendency among scholars to interpret the *Social Contract* as solely a treatise on political theory, thereby isolating it from the context of Rousseau's philosophy as a whole. These interpretations therefore provide a rather limited understanding of the purposes of Rousseau's political theory. For instance, Zev Trachtenberg has claimed that the *Social Contract* describes a society in which citizens have become sufficiently rational to recognise that it is in their interest not merely to attend to their own well-being, but rather to secure the common good of society as a whole. According to this reading, Rousseau's concept of the general will would then be an expression of this socialised rational self-

[577] Rousseau, *Émile* IV p.551: 'Mais considerez prémiérement, que voulant former l'homme de la nature il ne s'agit pas pour cela d'en faire un sauvage et de le reléguer au fond des bois, mai qu'enfermé dans le tourbillon social, il suffit qu'il ne s'y laisse entraîner ni par les passions ni par les opinions des hommes, qu'il voye par ses yeux, qu'il sente par son cœur, qu'aucune autorité ne le gouverne hors celle de sa propre raison.'

[578] Rousseau, *Émile* III p.484.

interest.[579] Another common reading, defended by Neuhouser among others, ascribes to Rousseau the view that the self-interested proclivity of civilised man may be overcome through public debate. By exchanging rational arguments, individual citizens would become inclined to transcend the self-interested bias of their own rational faculty and collectively attend to the public interest. The outcome of their discussion would then produce the general will.[580]

These readings of Rousseau's *Social Contract* are vulnerable to a number of objections.[581] Yet our present aim is not to provide a comprehensive interpretation of Rousseau's political theory or even of his concept of the general will. Instead, we will try to show that while the *Social Contract* may be read as primarily a treatise on the mechanics of popular sovereignty, Rousseau's political theory also suggests a number of measures by which the problems of modern society may be resolved by reconstructing the properties of life in the state of nature, thereby leading civilised man towards a more felicitous existence. To begin with, Rousseau maintains that the goal of the constitution of the *Social*

---

[579] Trachtenberg, *Making Citizens* 106: 'The intellectual advances that allow men to envision the future stimulate the awakening of socialized self-interest, so that men come to see the future advantages of sharing their efforts with others.' This reading essentially ascribes to Rousseau an account of the emergence of civilisation similar to that of Pufendorf, as well as Condillac and Helvétius, according to which individuals deduce by means of reasoning that it is in their long-term self-interest to attend to the common good rather than merely their own egocentric desires.

[580] Frederick Neuhouser, *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition* (Oxford 2011) 191-213.

[581] To Neuhouser's reading, we may object that the *Social Contract* II.iii p.61; IV.ii p.109 states that while citizens would decide political questions through voting, they should not be allowed to publicly discuss issues beforehand. Trachtenberg's interpretation is contradicted by Rousseau, *Geneva Manuscript* I.ii p.160: 'It is false that in the state of independence, reason leads us to cooperate for the common good out of a perception of our own interest. Far from being an alliance between private interest and the general good, they are mutually exclusive in the natural order of things, and social laws are a yoke that each wants to impose on the other without having to bear himself.' For an extensive reconstruction of the arguments for Rousseau's position that reason alone is not sufficient to constitute morality, see: Eve Grace, 'Built on Sand: Moral Law in Rousseau's Second Discourse', in: idem and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge 2013) 168-193, 190-192. For a promising attempt to resolve these problems, see David James, 'Rousseau on needs, language and pity: The limits of 'public reason'', *European Journal of Political Theory* 10 (2011) 372-393, who maintains that citizens of Rousseau's ideal polity would establish the general will not only by employing their reason, but also by attending to their sentiment of pity. Incited by the sentiment of pity, citizens of Rousseau's ideal polity will not merely attend to their self-interest, but rather employ their reason to imagine and sympathise with the concerns of their fellows, thereby attending to the good of society as a whole rather than only their self-interest.

Contract is to revive two conditions that had also been characteristic of the state of nature:

> If one seeks to define precisely what constitutes the greatest good of all, which ought to be the end of every system of legislation, one will find that it comes down to these two principal objects: *freedom* and *equality*.[582]

For the modern reader, Rousseau's emphasis on freedom and equality as the ultimate purpose of the state may appear commonplace. Yet this position was shared neither by Rousseau's predecessors, nor by many of his contemporaries.[583] In the remainder of this chapter, we will try to show that Rousseau's call for liberty and equality is not simply a fundamental axiom, but is in fact an attempt to relieve the predicament of civilised man by recreating some of the conditions of the state of nature in modern society.

To begin with, Rousseau's ideal state aims for equality among its citizens in order to recreate some of the conditions that accorded natural man a felicitous existence. Thus in the state of nature, humans had been equals simply because they had been solitary and without any durable moral relations. Rousseau states that in his ideal republic, this natural equality would be substituted by a moral equality that would negate the physical inequality that may nonetheless persist among its citizens. Rousseau thereby explicitly compares the equality of the state of nature with the moral equality ensured in his ideal society.[584] To be sure, this moral equality does not encompass total economic equality, even if Rousseau states that excessive disparities in wealth are to be avoided.[585] Rather, Rousseau conceives moral equality as equality of rights, which above all entails the abolition of privilege and class distinctions. While clearly an appeal to reform contemporary society, Rousseau's advocacy for moral equality is also motivated by his theory of human nature. Thus we have seen above that Rousseau thinks that inequality in society stimulates the development of *amour propre*, as citizens

---

[582] Rousseau, *Social Contract* II.xi p.75.

[583] Thus we have seen that for Hobbes, the purpose of the state is security and creating the conditions that allow for commodious living. According to Locke, the ultimate goal of the state is safeguarding the (property) rights of its citizens and resolving conflicts of rights as an independent arbiter.

[584] Rousseau, *Social Contract* I.ix p.58: 'rather than destroying natural equality, the fundamental compact on the contrary substitutes a moral and legitimate equality for whatever physical inequality nature may have placed between men.'

[585] Rousseau, *Social Contract* II.xi p.75

will covet the position and esteem of their superiors. At the same time, we have also seen that this inequality and associated development of *amour propre* will also impede the operation of man's conscience, as class distinctions will prevent citizens from identifying with each other. Consequently, citizens will tend to pursue their self-interest, thereby undermining the consensus on the common good required if self-government through the general will is to succeed. Conversely, moral equality is an essential requirement if citizens are expected to commiserate with their fellows as well as identify with the common good.[586]

The second objective of the political theory of the *Social Contract* is the recreation of liberty. Rousseau's advocacy for political liberty may simply originate in a conviction that despotism and slavery are incompatible with human nature.[587] Yet Rousseau also provides a more extensive argument why, by recreating the conditions of life in the state of nature, liberty is an essential prerequisite to human happiness. According to Rousseau man living in the state of nature has the right to anything he can get. As natural man is a solitary being, his liberty is therefore constrained solely by necessity.[588] By contrast, following the advent of sociability and institution of the state, the liberty of civilised man is constrained both by mutual dependence and laws devised by despotic government, thereby losing the capacity for self-determination. Rousseau thinks that this loss of self-determination is a crucial impediment to achieving felicity. As we have seen, Rousseau conceives felicity as a balance between desires and the ability to obtain desired objects. Accordingly, Rousseau thinks that our pursuit of happiness demands that potential obstacles to our volition must be removed as much as possible. Of course, Rousseau recognises that we cannot hope to escape the constraints of necessity and the laws of nature, which we may consider an unavoidable fact of life and which consequently do not infringe our liberty. But in society, it is not nature that poses the primary obstacles to our volition, but rather the opposing wills of fellow citizens as well as the authority of

[586] Rousseau, *Émile* IV p.503-505; 508-510.

[587] Rousseau, *On the Social Contract* I.iv p.50: 'To renounce one's freedom is to renounce one's status as a man, the rights of humanity and even its duties [...] Such a renunciation is incompatible with the nature of man, and taking away all his freedom of will is taking away all morality of his actions.' Andrew Levine, *The Politics of Autonomy: A Kantian Reading of Rousseau's* Social Contract (Amherst 1976) 14-15 indeed suggests that Rousseau's advocacy for liberty is simply the result of the axiom, unsupported by any further argumentation, that despotism and slavery are fundamentally opposed to human nature.

[588] Rousseau, *On the Social Contract* I.viii p.56.

the state. Sociability thereby introduces moral constraints that prevent us from doing what we want, thereby diminishing our happiness.

In response, Rousseau maintains that in order to recreate the self-determination of the state of nature, citizens themselves should be allowed to formulate the laws that would regulate their behaviour.[589] As citizens formulate their own laws, the constraints to their will are the consequence of their own volition. Accordingly, it would be incoherent if citizens experience these laws as inhibiting their liberty. Released from despotism, the volition of citizens will only be restrained by necessity, which is morally neutral, and by the legal restrictions they imposed on themselves. As the resulting laws are self-proclaimed by the general will of the people, they do not infringe the liberty of citizens.[590] The political liberty of the *Social Contract* thereby intends to revive the self-determination that had characterised life in the state of nature.[591] Like natural man, the citizens of the state designed in the *Social Contract* will no longer experience any moral constraints to their felicity. For this reason, Rousseau thinks that popular sovereignty is not only an essential requirement for liberty in society, but also a crucial prerequisite for the felicity of civilised man.[592]

Conclusion

Throughout our preceding discussion, we have seen that Rousseau's theory of human nature is clearly constructed upon empiricist principles. By placing his conception of the perfectibility of man in the context of his predecessors, we have shown that many of Rousseau's conclusions, in particular regarding his theory of human understanding, are adaptations or elaborations of the philosophies of Hobbes, Locke, and Condillac. Rather than a fundamental revision of their principles, Rousseau's most important contribution is the insight that the empiricist notion that man's ideas and intellectual faculties are acquired rather than innate entailed that human nature is neither universal nor static, but has

[589] Rousseau, *Émile* II p.311. Rousseau states that the liberty of self-government: 'réuniroit dans la République tous les avantages de l'état naturel à ceux de l'état civil.'

[590] Rousseau, *Social Contract* II.vi p.66. Accordingly Rousseau, *On the Social Contract* I.viii p.56 states that 'obedience to the law one has prescribed to oneself is freedom.'

[591] Judging from the substantial discussions on the Roman Republic in the *Social Contract*, Rousseau also sees his ideal polity as reviving the republican liberty of the ancient city-states. For a more extended discussion of Rousseau's political thought in the context of republicanism, see Annelien de Dijn, 'Rousseau and Republicanism', *Political Theory* 46 (2018) 59-80.

[592] Rousseau, *Émile* II p.311-312.

rather developed over the course of history. At the same time, we have also seen that Rousseau transcends the thought of his predecessors by attributing certain moral sentiments to human nature. These include the sentiment of pity and the development of *amour propre*. We have then established that Rousseau's theory of conscience may not be interpreted as a form of revived Platonism. While transcending his seventeenth-century predecessors, Rousseau's theory of human nature does accord with the intention of some of his contemporaries, including Hume and Diderot, to move beyond the reductionist tendencies of empiricism by ascribing to human nature a more extensive range of innate abilities and sentiments.

In the second part of the chapter, we have tried to elucidate the most important connections between Rousseau's theory of human nature and his moral and political philosophy. According to Rousseau, the concurrent emergence of society and development of human understanding by means of language have contributed to the misery of civilised man in several ways. Thus the development of reason has both inflated the passions and given rise to *amour propre*, which has in turn stifled the sentiment of pity. As a result, civilised man has become an ingenious yet covetous, envious and heartless being that spends its life searching for fleeting moments of pleasure, without ever attaining felicity. In response, Rousseau proposes to reform civilised man in the image of natural man, presenting two approaches by which this goal may be achieved. The first of these approaches is outlined in the *Émile*, and consists of a program of education that isolates its pupil from the most pernicious influences of modern society. Like natural man, the mature Émile will therefore become both self-sufficient and neither covetous nor envious.

Subsequently we have tried to show that the political theory of the *Social Contract*, which is often read in isolation, may also be conceived as a second solution to the problems resulting from the perfectibility of human nature. We have interpreted the institution of both liberty and equality by the constitution of the *Social Contract* as a means of amending the corruption of human nature that followed the emergence of society and invention of language. Thus political liberty allows citizens to formulate their own laws, thereby providing them with a limit to their desires that does not impede their felicity. Furthermore, moral equality mitigates the effects of *amour propre*, thereby encouraging the revival of the sentiment of pity in civilised man. As a result, citizens of the polity described in the *Social Contract* will become less covetous or envious. Less subject to

unattainable desires for recognition and material goods, these citizens will also achieve a greater measure of felicity. Yet the institution of liberty and equality will not merely transform them into more felicitous individuals. It will also turn the inhabitants of the state into more virtuous citizens who are more inclined to identify with the common good.

Conclusion

Following our investigation of the moral epistemology of Hobbes, Locke, Condillac, Helvétius, and Rousseau, we may now outline a number of general conclusions. To begin with, our discussion has tried to showcase that these early-modern philosophers did not consider epistemology and ethics as separate fields of inquiry. Rather, we have seen that their epistemological inquiries are concerned not only with describing the means by which we may acquire knowledge of empirical objects, but also raise the question how humans may use their reason to understand morality. At the same time, we have seen that the moral philosophies by which they attempt to resolve this ethical question are generally accommodated to their more general theories of knowledge and human understanding. Accordingly, our investigation has tried to demonstrate that the epistemology of Hobbes, Locke, Condillac, Helvétius, and Rousseau serves as a valuable context to any attempt at a comprehensive interpretation of their respective moral philosophies. In our remaining paragraphs, we will briefly reiterate the questions of moral epistemology that confronted these thinkers, and outline their answers as we have reconstructed them in our interpretation.

To begin with, we have seen that all five philosophers agree that there is a universal standard of right and wrong that may be formulated as natural law. The starting point of our discussion has therefore been to inquire by what method these philosophers think that human beings are able to grasp this universal moral standard. Yet throughout our discussion we have seen that answering this question is less straightforward than may at first be expected. It turns out that for all five philosophers, their account of the foundations of morality is complicated by at least two important properties of their epistemology. The first of these is the tendency, common to all five thinkers, to explain human cognition while ascribing to the mind as few innate properties as possible. Thus following Locke, they all dismiss the notion that human beings are endowed with innate ideas that may serve as the foundation for moral reasoning. Yet these philosophers also reject the conception of reason as an autonomous faculty innate to human nature. Rather, they present human cognition as nothing but a process, set in motion by the passions, of comparing and connecting ideas – the cognitive units that provide the contents of our thoughts that may ultimately be traced back to experience. This account of human cognition cannot easily be

reconciled with any attempt to present the foundations of morality as innate in the mind itself.

Secondly, the moral epistemology of Hobbes, Locke, Condillac, Helvétius, and Rousseau is further confounded by their empiricism and their representationalist conceptions of knowledge. Thus we have seen that all these thinkers maintain that knowledge of empirical reality may only be acquired by means of experience. Accordingly, the real essences of empirical objects will remain hidden from us, as we are solely able to experience their sensory properties. Nonetheless, these philosophers think that humans may acquire at least highly probable knowledge of substances if our ideas are made to represent the empirical objects we experience. Even if representationalist conceptions of knowledge like those found in Locke and his followers are nowadays widely contested, these philosophers provided an account of the origins of knowledge of empirical objects that was at least reasonably convincing to their contemporaries.[593] Yet we have seen that by itself, this representationalist conception of knowledge cannot explain how we may acquire proper ideas of morality. Unlike ideas of substances, moral concepts like 'justice' or 'liberty' do not represent any concrete objects of empirical reality and accordingly cannot simply be experienced by means of the senses. The empiricism of the five philosophers discussed in this thesis is thus not only opposed to the theory that a rational understanding of morality would be innate to human nature, but also confound any theory that morality would be a property of the external world.

Yet if our understanding of morality is neither innate to human nature, nor acquired by means of experience, how do these five philosophers think that humans may grasp the universal morality of natural law? Throughout our discussion we have seen that these authors themselves do not always provide a straightforward answer to this question. Accordingly, we have found that modern scholars have reconstructed their answers in variety of diverging interpretations. For instance, we have seen that interpreters of both Hobbes and Rousseau, unable to reconstruct a coherent account of moral epistemology from their writings, have maintained that when these thinkers claim that natural law is accessible to reason, they simply assume that this normative reason is an innate property of human nature. These interpreters thereby claim that while Hobbes and Rousseau advance an empiricist account of the origin of knowledge,

[593] E.g. Rorty, *Philosophy and the Mirror of Nature* 230-256.

they nonetheless retain the rationalist assumption that the ability for normative judgment is innate to human nature. Yet throughout our own investigation, we have seen that Hobbes and Rousseau describe reason as simply the process of comparing and relating ideas present in the mind. Following our discussion of Hobbes's and Rousseau's epistemologies, we have therefore concluded that these thinkers could not have simply assumed that normative reason is somehow innate to human nature.

Secondly, we have encountered in the scholarship on Hobbes and Locke the interpretation that our understanding of morality could be derived from the conventional signification of moral terms. This interpretation lends credibility from the fact that Hobbes and Locke, as well as their followers, all claimed that the more advanced cognitive abilities of the human mind are developed by means of language. In particular, the thinkers discussed in this thesis agree that the composition of abstract or universal concepts is only made possible by the use of words. They thereby explain why humans may acquire advanced forms of cognition not accessible to animals, but without supposing that the human mind is endowed with any additional innate cognitive faculties. Yet as we have seen, this explanation also invites the suggestion that words are not merely cognitive tools, but that their signification may also influence the ideas in the minds of individuals. Thus we have seen that on the one hand, the five philosophers claim that words are indispensible for the composition of complex ideas. But on the other hand, we have also seen that these thinkers became increasingly aware of the fact that words are not merely private marks to ideas but also allow us to communicate our ideas to others. The signification of words is therefore circumscribed by linguistic conventions. As we are compelled to adhere to these conventions in order to be understood, this suggests that the composition of our ideas is influenced or possibly even determined by the standard of common use. This suggestion is particularly relevant in the case of moral concepts, as these are abstract concepts that do not represent any concrete objects of experience, and may only be rendered in linguistic terms.

Accordingly, we have seen that a number of scholars have ascribed to Hobbes and Locke the view that the moral concepts retained by individuals are regulated by the conventional signification of their moral terminology. Yet our own discussion has established that while Hobbes's and Locke's theories of signification may appear to suggest this conclusion, neither thinker explicitly endorses it. Rather, we have seen that the theory that the conventional

signification of terms determines the composition of ideas in the minds of individuals is only fully developed in the philosophy of Condillac. As he eradicates any remaining assumption that the mind is an autonomous entity, Condillac claims that both our ideas and our cognitive faculties are acquired through experience and organised by means of language. Drawing a much more explicit connection between language and ideas, Condillac maintains that the revision of the latter requires the wholesale reform of the former. Accordingly, we have seen that Condillac concludes that moral progress requires the amendment of the conventional signification of moral terminology. In subsequent chapters, we have then seen that Condillac's theory on the relation between language and ideas is not only a development of suggestions implicit in the philosophies of Hobbes and Locke, but also an important influence on both Helvétius and Rousseau. Thus we have seen that Helvétius adopts Condillac's conception of the mind as developed by means of experience and language, and likewise recommends the reform of the conventional signification of moral terms. Furthermore, our last chapter has shown that Condillac's epistemology and philosophy of language also provided the main source of inspiration for Rousseau's conception of human nature as perfectible.

But if neither innate faculties nor linguistic conventions are the source of our rational understanding of morality, how does normative reasoning operate according to the five philosophers discussed in this thesis? In our discussion, we have presented their accounts of practical deliberation as the key to providing an answer to this question. To begin with, we have seen that these thinkers proceed from the assumption that all human thoughts and actions are ultimately motivated by a desire for pleasure and aversion of pain. Accordingly, some modern readers have concluded from this assumption that according to these thinkers, humans are necessarily determined by their passions, while their reason is merely instrumental and relegated to finding the means to whatever they desire. Yet in particular in our discussions of Hobbes's and Rousseau's accounts of practical deliberation, we have demonstrated that apart from a few corporeal appetites like hunger or thirst, they do not consider the passions as mindless urges. Rather, the passions originate in mental discourse and could well be described as beliefs about the foreseeable consequences of a certain object or action. According to our reading, reason is therefore not merely the 'slave of the passions,' limited to finding the means to whatever we desire. Rather, as it may foresee and weigh the consequences of our actions, reason in fact plays an

additional role in practical deliberation by assessing whether objects are in fact truly desirable.

Subsequently, we have established that for these thinkers apart from Rousseau, reason participates in normative judgment in a similar way as it engages in practical deliberation. First of all, we have seen that these philosophers assume that the primary aim of morality is to guide humans towards felicity. Yet as they claim that attaining pleasure and avoiding pain is the ultimate aim of all human thought and action, the philosophers discussed in this thesis, again with the exception of Rousseau, conclude that this felicity consists in nothing but the durable satisfaction of our passions. They therefore conceive morality as ultimately nothing but a prudential rule that would guide humans towards this goal. The way in which reason participates in subjective normative judgment is similar to its role in practical deliberation. As reason may not only suggest the means to the objects of our desire, but also foresee the consequences of the pursuit of our passions, it may formulate general prudential rules that would commend actions with generally pleasurable consequences, thereby leading individuals towards felicity. Accordingly, we have classified the account of moral reasoning of these four thinkers as forms of consequentialism, according to which normative judgment relies on foreseeing the consequences of our actions.

Yet while this account of subjective moral judgment describes how humans may use their reason to pursue their felicity from an individual perspective, it does not yet explain how they may ascertain the moral guidelines that would also ensure their happiness in society. Even if they do not consider humans as naturally social beings, the philosophers discussed above all concur that humans generally do live together in society. Forced to share limited resources, humans are bound to come into conflict with each other when they pursue their passions without regard for the interests of their fellows. As such conflicts tend to be mutually destructive, prudential reasoning dictates that humans living in society anticipate the actions of their fellows. In addition, individuals should try to contemplate the consequences of their own pursuits for those around them. For humans living in society, morality is thus not merely a prudential guide towards the satisfaction of their own passions, but rather a communal rule that, if universally upheld, will diffuse conflicts of interest and provide the ideal conditions in which citizens may pursue their felicity. Yet once we realise that the aim of morality is not merely the felicity of the individual, but

also the good of society as a whole, the question arises how individuals may employ their reason to ascertain the most prudential rules that would not merely promote their self-interest, but also the common good.

Throughout this thesis, we have seen that our five philosophers have each formulated a different answer to this question. In our first chapter, we have shown that Hobbes aims to derive a universal morality from a substantive conception of human nature. Thus we have seen that Hobbes prefaces his moral and political philosophy with an account of the development of human cognition, as well as a detailed investigation of the mechanism by which the passions provide the motivation for all of our actions. Based on this theoretical account of human nature, Hobbes subsequently demonstrates in his famous thought experiment that individuals living in the state of nature are destined to a gruesome existence. Unchecked by any centralised authority, humans are prone to pursue their selfish passions to the detriment of others, leading first to diffidence and ultimately to a state of war. Yet following his investigation of human nature, Hobbes also identifies the passions that dispose individuals towards a peaceful and more pleasant existence. Thus Hobbes concludes that in particular the desire for self-preservation, as well as the inclination for commodious living and the aspiration to acquire knowledge, disposes individuals towards following the laws of nature and subjecting to centralised authority. We have then claimed that the reasoning leading Hobbes to this conclusion employs an approach similar to that of practical deliberation. Scrutinising the various passions prevalent in human nature in view of their consequences, Hobbes has employed reason to draw normative conclusions by foreseeing which desires should be pursued as they direct humans towards felicity, and which passions should ignored due to their propensity to cause misery. His natural laws are then the prudential guidelines towards satisfying the one passion that has favourable consequences under all circumstances – the desire for self-preservation.

In our next chapter, we have seen that John Locke is prevented from adopting Hobbes's approach to moral epistemology as it conflicts with at least two central tenets of his philosophy. To begin with, Locke's epistemology introduces a strict distinction between ideas of substances acquired by means of experience, and ideas of modes such as moral concepts that are constructed by the mind at will. Even if Locke does not spell out the consequences of this position for morality with the same transparency as his successor David Hume, we have nonetheless seen that Locke to identifies a dichotomy between facts and

values in his epistemology. In line with his own prohibition against the use of ideas of substances in moral reasoning, Locke then explicitly rejects the practice of adopting a substantive conception of human nature as the foundation for morality. For Locke, the conception of human nature to be used in moral reasoning describes humans solely as individuals motivated by a desire for pleasure and aversion of pain, who may employ their capacity for reasoning to achieve this aim. In isolation, this conception of human nature would provide grounds to interpret Locke as a consequentialist and perhaps (proto-)utilitarian moralist. Yet we have seen that this reading, proposed by several modern interpreters, is impeded by Locke's adherence to the notion that morality is not simply a prudential guideline towards felicity during life on earth, but rather a divine decree enforced by sanctions and rewards administered in the afterlife. We have concluded that together with his injunction against deriving normative conclusions from a substantive conception of human nature, Locke's adherence to the notion of morality as a divine decree prevents us from reconstructing an entirely coherent interpretation of Locke's moral epistemology.

In the following chapter, we have seen that Condillac derives his theory of human nature almost entirely from Locke. Thus Locke's conception of man as a being motivated by the passions and guided by reason also provides the basis for Condillac's moral philosophy. Yet we have observed that following this account of human motivation, Condillac is hardly concerned with any ulterior questions of moral epistemology. Thus Condillac simply thinks that if humans employ their reason in order to foresee the consequences of their actions, they would realise that it would be in their long-term self-interests to become social beings and act in accordance with the common interest. At the same time, we have seen that Condillac develops Locke's reflections on the role of language in cognition into a comprehensive theory of linguistic determinism. More preoccupied with the role of language in human cognition, Condillac thus concludes that moral progress does not necessarily require a novel approach to moral epistemology, but rather the reform of the conventional signification of moral terminology.

Subsequently, we have seen that while Helvétius also builds upon the empiricist conception of human nature proposed by Locke and further developed by Condillac, he presents an account of human cognition that severely simplifies the theories of his predecessors. We have therefore concluded that the main purpose of Helvétius's philosophical inquiries may not have been to present a

credible and coherent theory, but rather to provide abstract argumentation to support his call to moral and political reform. Nonetheless, the reductionism of Helvétius's moral philosophy highlights two important properties of the empiricist moral theories we have investigated in this thesis. Firstly, by claiming that the mind is nothing but a collection of ideas endowed with sensibility, Helvétius draws attention to the potentially mechanistic character of empiricist accounts of human understanding that claim that all ideas are derived from sensation and all actions are occasioned by the passions. Earlier empiricists had tried to avoid or mitigate the consequences of determinism, either by tacitly assuming that the mind is an autonomous entity separate from its ideas, or by claiming that language allows human understanding to reason independent from the immediate stimuli of sensation. Yet as he largely ignores Condillac's account of the role of language in cognition, and takes the empiricist tendency to ascribe to human nature as few innate properties as possible to its logical conclusion, Helvétius presents the mind as an entirely passive entity, wholly determined by its passions and the external stimuli of sensation. Following the critiques of contemporaries like Diderot and Rousseau, we have concluded that this theory can no longer be taken seriously as a comprehensive account of human nature.

Secondly, Helvétius's moral philosophy takes the crucial step of transforming the qualified consequentialism of his predecessors into a genuine theory of moral utilitarianism. While Hobbes, Locke, and Condillac all retain consequentialist notions of normative judgment, they do not conclude that the purpose of morality is therefore simply the greatest happiness of the greatest number. Thus we have seen that following his consequentialist account of normative judgment, Hobbes conducts a scrutiny of the passions prevalent in human nature and identifies the desires likely to have the most agreeable consequences, concluding that all rational individuals would agree that the desire for self-preservation should take precedence above all other passions. Accordingly, the aim of Hobbes's natural laws is not simply to maximise pleasure and minimise pain, but rather to guide humans towards their self-preservation. Subsequently, we have seen that Locke's consequentialist account of normative judgment is curtailed by his belief in Divine retribution. For Locke, the aim of morality is thus not simply the maximisation of pleasure during life on earth, but rather to direct us towards attaining eternal bliss in the afterlife. Finally, Condillac presents the argument, common among many eighteenth-century moralists, that the passions generally dispose human beings to cooperate, as they

realise that they are better able to satisfy their needs when united in society. Condillac's consequentialism thereby merely provides him with an explanation of human sociability. Yet Helvétius, true to his generally less nuanced approach to philosophy, simply concludes from the position that all human actions are motivated by the passions that the aim of morality and the state should be to ensure felicity by maximising pleasure and minimising pain. By uniting a consequentialist theory of normative judgment with the principle of the greatest happiness of the greatest number, Helvétius presents one of the first consistent formulations of modern utilitarianism.

In our final chapter we have discussed the moral philosophy of Rousseau, arguably the most influential early-modern critic of the consequentialist moral theories proposed by the other thinkers featured in this thesis. Yet even if Rousseau's conclusions diverge sharply from those of his predecessors, we have tried to show that his epistemology, philosophy of language and account of practical deliberation are nonetheless adaptations of the theories of Hobbes, Locke, and Condillac. Rather than adopting fundamentally different preconceptions, Rousseau reassesses their theories on a few crucial points, leading him to conclusions notably opposed to those of his predecessors. Thus we have seen that Rousseau's conception of human nature as perfectible is developed from Locke's and above all Condillac's theories on language and cognition. Adopting Locke's notion that all ideas in the mind are acquired by means of experience, as well as Condillac's theory that human cognition is only fully developed with the help of language, Rousseau concludes that the human mind is not only a *tabula rasa* at birth, but that human nature would have been a blank slate at the dawn of humankind as well. Rousseau therefore presents human nature as malleable rather than universal. Accordingly, Rousseau dismisses the approach, attempted by Hobbes for instance, of adopting a conception of human nature as the foundation for a universal theory of morality.

Furthermore, we have also seen that Rousseau's revision of Hobbes's account of practical deliberation leads him to a rejection of the notion of morality as a prudential guideline towards felicity, conceived as nothing but the sustained satisfaction of our desires. Rousseau maintains that felicity cannot be achieved simply by expanding our foresight, as the development of reason will also increase the objects of desire that will disturb our contentment. Alternatively, Rousseau maintains that felicity can be attained only by achieving a balance between the passions and the means to satisfy our desires. We have then

identified two methods by which Rousseau thinks that this balance between passions and aptitude may be achieved. The first method, outlined in the *Émile*, consists of a program of education in which a pupil is brought up in isolation to become an autonomous individual unperturbed by the seductions of modern society. We have subsequently maintained that the political theory of the *Social Contract* provides a second approach towards this aim by allowing citizens to formulate their own laws. As the sole impediments to the volition of citizens will be either of necessity, or self-declared by means of the general will, their desires will no longer extend beyond those things that nature and the law have assigned to them. Rousseau's revised conception of felicity as a balance between the passions and the means to satisfy desires thereby provides a potent argument in favour of moral autonomy and self-governance. Ultimately on the basis of only a slight amendment of Hobbes's account of practical deliberation, Rousseau has therefore formulated a powerful critique of all consequentialist and utilitarian moral theories that assume that felicity is nothing but the satisfaction of our desires, and that morality is merely a prudential rule guiding us towards this goal.

Bibliography

Aarsleff, Hans, 'Condillac's Speechless Statue', in: idem, *From Locke to Saussure* (Athlone: London 1982) 210-224.

Aarsleff, Hans, 'introduction' in: Condillac, *Essay on the Origin of Human Knowledge* ed. and trans. by Aarsleff (Cambridge University Press: Cambridge 2001).

Aarsleff, Hans, 'Pufendorf and Condillac on Law and Language', *Journal on the Philosophy of History* 5 (2011) 308-321.

Aarsleff, Hans, 'The State of Nature and the Nature of Man in Locke', in: John W. Yolton (ed.), *John Locke: Problems and Perspectives* (Cambridge University Press: Cambridge 1969) 99-136.

Aarsleff, Hans, 'The Tradition of Condillac: The Problem of the Origin of Language in the Eighteenth Century and the Debate in the Berlin Academy before Herder', in: idem, *From Locke to Saussure, Essays on the Study of Language and Intellectual History* (Athlone: London 1982) 146-209.

Abizadeh, Arash, 'Hobbes on Mind: Practical Deliberation, Reasoning, and Language', *Journal of the History of Philosophy* 55 (2017) p.1-34.

Adams, Marcus, 'Hobbes, Definitions and Simplest Conceptions', *Hobbes Studies* 27 (2014) 35-60.

Ashworth, Jennifer, '"Do Words Signify Ideas or Things?" The Scholastic Sources of Locke's Theory of Language', *Journal of the History of Philosophy* 19 (1981) 299-326.

Atherton, Margaret, 'Locke on Essences and Classification', in: Lex Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* (Cambridge University Press: Cambridge 2007) 258-285.

Audidière, Sophie, 'Why do Helvétius's writings matter? Rousseau's *Notes sur De l'esprit*', *British Journal for the History of Philosophy* 24 (2016) 983-1001.

Auroux, Sylvain, 'Empirisme et théorie linguistique chez Condillac', in : Sgard (ed.), *Condillac et les problèmes du langage* (Slatkine: Geneva 1982) 177-219

Ayers, Michael, *Locke I: Epistemology* (Routledge: London and New York 1991).

Ayers, Michael, *Locke II: Ontology* (Routledge: London and New York 1991).

Barnouw, Jeffrey, 'Reason as Reckoning: Hobbes'ss Natural Law as Right Reason', *Hobbes Studies* 21 (2008) 38-62.

Berkeley, *Principles of Human Knowledge* ed. by Howard Robinson (Oxford University Press: Oxford 1996).

Blau, Adrian, 'Reason, Deliberation, and the Passions', in: A.P. Martinich and Kinch Hoekstra (eds.), *The Oxford Handbook of Hobbes* (Oxford University Press: Oxford 2016) 195-217.

Bluhm, William T., 'Locke's Idea of God: Rational Truth of Political Myth?', *The Journal of Politics* 42 (1980) 414-438.

Bolton, Martha Brandt, 'The Taxonomy of Ideas in Locke's Essay', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 67-100

Botros, Sophie, *Hume, Reason and Morality: A Legacy of Contradiction* (Routledge: London and New York 2006).

Brooke, Christopher, *Philosophic Pride: Stoicism and Political Thought from Lipsius to Rousseau* (Princeton University Press: Princeton 2012).

Burson, Jeffrey D., *The Rise and Fall of Theological Enlightenment: Jean-Martin de Prades and Ideological Polarization in Eighteenth-Century France* (Notre Dame Press: Notre Dame 2010).

Byron, Michael, 'Hobbes's Confounding Foole', in: S.A. Lloyd (ed.), *Interpreting Hobbes's Political Philosophy* (Cambridge University Press: Cambridge 2019) 206-222.

Callaghan, G.K., 'Nominalism, Abstraction, and Generality in Hobbes', *History of Philosophy Quarterly* 18 (2001) p.37-55.

Cassirer, Ernst, *The Philosophy of the Enlightenment* (Princeton University Press: Princeton 2009 [1951]).

Cassirer, Ernst, *The Question of Jean-Jacques Rousseau* ed. and trans. by Peter Gay (Indiana University Press: Bloomington 1954).

Charvet, John, *The Social Problem in the Philosophy of Rousseau* (Cambridge University Press: Cambridge 1974).

Coli, Daniela, 'Hobbes's Revolution', in: Victoria Kahn et al. (eds.), *Politics and the Passions 1500-1850* (Princeton University Press: Princeton 2006)73-92.

Colman, John, *John Locke's Moral Philosophy* (Edinburgh University Press: Edinburgh 1983).

Condillac, 'Extrait raisonnée', in: idem, *Traité des sensations augmenté de l'extrait raisonnée* (Fayard: Paris 1984).

Condillac, *Essay on the Origin of Human Knowledge* ed. and trans. by Hans Aarsleff (Cambridge: Cambridge University Press 2001)

Condillac, *Grammaire* in: *Oeuvres Philosophiques* ed. by Georges le Roy (Paris: Presses Universitaires de France 1947-1951).

Condillac, *Logic* in: *Philosophical Writings of Etienne Bonnot, Abbé de Condillac* trans. by Franklin Philip and Harlan Lane (Lawrence Erlbaum: Hillsdale 1982).

Condillac, *Logic* in: *Philosophical Writings* trans. by Philip and Lane (Lawrence Erlbaum: Hillsdale 1982).

Condillac, *Traité des animaux* ed. by M. Malherbe (Vrin: Paris 2004).

Condillac, *Traité des animaux* ed. by M. Malherbe (Vrin: Paris 2004).

Condillac, *Treatise on Sensations* in: *Philosophical Writings of Etienne Bonnot, Abbé de Condillac* trans. by Philip and Lane (Lawrence Erlbaum: Hillsdale 1982).

Condillac, *Treatise on Sensations* in: *Philosophical Writings* trans. by Philip and Lane.

Cranston, Maurice, 'Rousseau's Theory of Liberty', in: Robert Wokler (ed.), *Rousseau and Liberty* (Manchester University Press : Manchester 1995).

d'Alembert, *Essai sur les éléments de philosophie* ed. by Richard Schwab (Georg Olms: Hildesheim 1965).

d'Holbach, *Système de la nature ou les lois du Monde Physique et du Monde Moral* ed. by Jean-Pierre Jackson (Coda: Paris 2008 [1770]).

Damrosch, Leo, *Jean-Jacques Rousseau: Restless Genius* (Houghton Mifflin: Boston 2005).

Darwall, Stephen, 'Normativity and Projection in Hobbes's Leviathan', *The Philosophical Review* 109 (2000) p.313-347.

Darwall, Stephen, 'The Foundations of Morality: Virtue, Law and Obligation', in: Donald Rutherford (ed.), *The Cambridge Companion to Early Modern Philosophy* (Cambridge University Press: Cambridge 2006) 221-247.

Dawson, Hannah, 'Locke on Private Language', *British Journal for the History of Philosophy* 11 (2003) 609-638.

Dawson, Hannah, *Locke, Language and Early-Modern Philosophy* (Cambridge University Press: Cambridge 2007).

Deigh, John, 'Reason and Ethics in Hobbes's Leviathan', *Journal for the History of Philosophy* 34 (1996) 33-60.

Delaney, James, *Rousseau and the Ethics of Virtue* (Continuum: London 2006).

Dent, N.J.H., and O'Hagan, Timothy, 'Rousseau on "Amour-Propre"', *Proceedings of the Aristotelian Society* 99 (1999) 91-107.

Derathé, Robert, *Jean-Jacques Rousseau et le science politique de son temps* (Vrin: Paris 1988).

Diderot, *D'Alembert's Dream* trans. by Leonard Tancock (Penguin: London 1966).

Diderot, *Rameau's Nephew* trans. by Leonard Tancock (Penguin: London 1966).

Diderot, *Refutation d'Helvétius* in: *Oeuvres completes* XXIV ed. by H. Dieckmann and J. Varloot (Hermann: Paris 2004).

Diderot, *Supplément au voyage de Bougainville* in: John Mason and Robert Wokler (eds. and trans.), *Political Writings* (Cambridge University Press: Cambridge 1992).

Dijn, Annelien de, 'Rousseau and Republicanism', *Political Theory* 46 (2018) 59-80.

Dobel, J. Patrick, 'The Role of Language in Rousseau's Political Thought', *Polity* 18 (1986) 638-658.

Douglass, Robin, *Hobbes and Rousseau: Nature, Free Will, and the Passions* (Oxford University Press: Oxford 2015).

Duchesneau, François, 'Condillac et le principe de liaison des idées', *Revue de Métaphysique et de Morale* 1 (1999) 53-79.

Duncan, John, 'Perfectibility, Chance, and the Mechanism of Desire Multiplication in Rousseau's Discourse on Inequality', in: Mark Blackwell et al. (eds.), *Rousseau and Desire* (Toronto: University of Toronto Press 2009) 17-45.

Duncan, Stewart, 'Hobbes, Signification, and Insignificant Names', *Hobbes Studies* 24 (2011) 158-178.

Force, Pierre, 'Helvétius as an Epicurean Political Theorist' in: Neven Leddy and Avi S. Lifschitz (eds.), *Epicurus in the Enlightenment* (Voltaire Foundation: Oxford 2009).

Force, Pierre, *Self-Interest before Adam Smith, A Geneology of Economic Science* (Cambridge University Press: Cambridge 2003).

Forde, Steven, '"Mixed Modes" in John Locke's Moral and Political Philosophy', *The Review of Politics* 73 (2011) 581-608.

Forde, Steven, 'Natural Law, Theology, and Morality in Locke' *Journal of Political Science* 45 (2001) 396-409.

Forde, Steven, *Locke, Science and Politics* (Cambridge University Press: Cambridge 2014).

Frazer, Michael, *The Enlightenment of Sympathy: Justice and the Moral Sentiments in the Eighteenth Century and Today* (Oxford University Press: Oxford 2010).

Gaukroger, Stephen, *The Collapse of Mechanism and the Rise of Sensibility: Science and the Shaping of Modernity 1680-1760* (Oxford University Press: Oxford 2010).

Gauthier, David, 'Hobbes on Demonstration and Construction', *Journal of the History of Philosophy* 35 (1997) 509-521.

Gauthier, David, *The Logic of the Leviathan: The Moral and Political Theory of Thomas Hobbes* (Clarendon Press: Oxford 1969).

Gert, Bernard, 'Hobbes on language, metaphysics, and epistemology', *Hobbes Studies* 14 (2001) 40-58.

Gert, Bernard, 'Hobbes on Reason', *Pacific Philosophical Quarterly* 82 (2001) 243-257.

Gert, Bernard, *Hobbes, Prince of Peace* (Polity: Cambridge 2010).

Gillespie, Michael Allen, *The Theological Origins of Modernity* (University of Chicago Press: Chicago 2008).

Grace, Eve, 'Built on Sand: Moral Law in Rousseau's Second Discourse', in: idem and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge University Press: Cambridge 2013) 168-193.

Grace, Eve, 'The Restlessness of "Being": Rousseau's Protean Sentiment of Existence', *History of European Ideas* 27 (2001) 133-151.

Grotius, *On the Law of War and Peace* ed by Stephen Neff (Cambridge University Press: Cambridge 2012).

Haakonssen, Knud, *Natural Law and Moral Philosophy: from Grotius to the Scottish Enlightenment* (Cambridge University Press: Cambridge 1996).

Haakonssen, Knud, 'Natural Law without Metaphysics: a Protestant Tradition', in: Ana Martinez Gonzalez (ed.), *Contemporary Perspectives on Natural Law* (Routledge: London 2008) 67-85.

Haakonssen, Knud, 'The History of Eighteenth-Century Philosophy: History or Philosophy?', in: idem (ed.), *The Cambridge History of Eighteenth-*

Century Philosophy (Cambridge University Press: Cambridge 2006) I 3-25, 14.

Hall, John, *Rousseau: An Introduction to his Political Philosophy* (Schenkman: Cambridge Mass 1973).

Hanley, Ryan Patrick, 'Rousseau's Virtue Epistemology', *Journal of the History of Philosophy* 50 (2012) 239-263.

Harari, Yuval Noah, *Sapiens, A Brief History of Humankind* (Vintage: London 2011).

Helvétius, *De l'esprit* ed. by Jacques Moutaux (Fayard: 1988).

Helvétius, *De l'homme* ed. by Jacques Moutaux (Fayard: 1989).

Heydt, Colin, 'Utilitarianism before Bentham', in: Ben Eggleston and Dale E. Miller (eds.), *The Cambridge Companion to Utilitarianism* (Cambridge University Press: Cambridge 2014) 16-37.

Hill, Benjamin, 'Locke on the Names of Modes', in: Margaret Cameron and Robert J. Stainton (eds.), Li*nguistic Content: New Essays on the History of Philosophy of Language* (Oxford University Press: Oxford 2015) 176-200.

Hirschmann, Albert, *The Passions and the Interests: Political Arguments for Capitalism before its Triumph* (Princeton University Press: Princeton 1977).

Hobbes, *De cive* ed. Howard Warrender (Oxford: Oxford University Press 1983)

Hobbes, *De corpore* ed. by William Molesworth (Bohn: London 1839).

Hobbes, *Leviathan* ed. by Edwin Curley (Hackett: Indianapolis and Cambridge 1994).

Hobbes, *Questions concerning Liberty, Necessity and Chance* XXVI ed. William Molesworth (London: Bohn 1841)

Hochstrasser, Tim, 'Conscience and Reason: The Natural Law Theory of Jean Barbeyrac', *The Historical Journal* 36 (1993) 289-308.

Hochstrasser, Tim, *Natural Law Theories in the Early Enlightenment* (Cambridge: Cambridge University Press 2000).

Hoekstra, Kinch, 'Hobbes on Law, Nature, and Reason', *Journal of the History of Philosophy* 41 (2003) 111-120.

Hruscka, Joachim, 'The Greatest Happiness Principle and Other Early German Anticipations of Utilitarian Theory', *Utilitas* 3 (1991) 165-177.

Hulliung, Mark, *The Autocritique of Enlightenment: Rousseau and the Philosophes* (Transaction: New Brunswick 2014 [1994]).

Hume, *A Treatise of Human Nature* ed. by David Fate Norton and Mary J. Norton (Oxford University Press: Oxford 2000).

Hume, *An Enquiry concerning Human Understanding* ed. by Peter Nidditch (Clarendon Press: Oxford 1975).

Hume, *An Enquiry concerning the Principles of Morals* ed. by Tom L. Beauchamp (Oxford University Press: Oxford 1998)

Hundert, E.J., *The Enlightenment's Fable: Bernard Mandeville and the Discovery of Society* (Cambridge University Press: Cambridge 2005).

Israel, Jonathan, *Democratic Enlightenment: Philosophy, Revolution, and Human Rights 1750-1790* (Oxford University Press: Oxford 2011).

Israel, Jonathan, E*nlightenment Contested: Philosophy, Modernity, and the Emancipation of Man 1670-1752* (Oxford University Press: Oxford 2006).

Israel, Jonathan, *Radical Enlightenment: Philosophy and the Making of Modernity 1650-1750* (Oxford University Press: Oxford 2001).

Israelsen, Andrew, 'God, Mixed Modes and Natural Law: An Intellectualist Interpretation of Locke's Moral Philosophy', *British Journal for the History of Philosophy* 21 (2013) 1111-1132.

Jacquette, Dale, 'Condillac's Analytic Dilemma', *History of Philosophy Quarterly* 32 (2015) 141-159.

James, David, 'Rousseau on Needs, Language and Pity: The Limits of 'Public Reason'', *European Journal of Political Theory* 10 (2011) 372-393.

James, Susan, *Passion and Action: the Emotions in Seventeenth Century Philosophy* (Oxford University Press: Oxford 1999).

Jimack, Peter, *La Genèse et la rédaction de l'Emile de J.-J.Rousseau* (Droz: Geneva 1960).

Jolley, Nicholas, *Locke: his Philosophical Thought* (Oxford University Press: Oxford 1999).

Joly, André, 'De la théorie du langage à l'analyse d'une langue', in: Sgard (ed.), *Condillac et les problèmes du langage* 243-256.

Kavka, Gregory, *Hobbesian Moral and Political Theory* (Princeton University Press: Princeton 1987).

Knight, Isabel, *The Geometric Spirit: The Abbé de Condillac and the French Enlightenment* (Yale University Press: New Haven 1968).

Kohler, Timothy, et al., 'Greater post-Neolithic wealth disparities in Eurasia than in North America and Mesoamerica', *Nature* 551 (2017) 619-622.

Kretzmann, Norman, 'The Main Thesis of Locke's Semantic Theory', *The Philosophical Review* 77 (1968) 175-196.

la Vopa, Anthony, 'A New Intellectual History? Jonathan Israel's Enlightenment', *The Historical Journal* 52 (2009) 717-738.

Ladd, Jr., Everett C., 'Helvétius and D'Holbach: "La Moralisation de la Politique" *Journal of the History of Ideas* 23 (1962) 221-238.

LeBuffe, Michael, 'Hobbes's Reply to the Fool', *Philosophy Compass* 2 (2007) 31-45.

Leibniz, *New Essays on Human Understanding* ed. by Peter Remnant and Jonathan Bennett (Cambridge University Press: Cambridge 1996)

Leisinger, Matthew, 'Locke's arguments against freedom to will', *British Journal for the History of Philosophy* 25 (2017) 642-662.

Lenz, Martin, *Lockes Sprachkonzeption* (De Gruyter: Berlin and New York 2010).

Lenz, Martin, 'Locke as a Social Externalist', *Proceedings of the British Academy* 189 (2013) 53-67.

Lenz, Martin, 'Locke's Lifeworld: The Teleological Role of Secondary Qualities', *History of Philosophy Quarterly* 36 (2019) 39-59.

Levine, Andrew, *The Politics of Autonomy: A Kantian Reading of Rousseau's Social Contract* (University of Massachusetts Press: Amherst 1976).

Levitin, Dmitri, 'Reconsidering John Sergeant's Attacks on Locke's *Essay*,' *Intellectual History Review* 20 (2010) 457-477.

Lifschitz, Avi, 'How to do Things with Signs: Rousseau's Ancient Performative Idiom', *History of Political Thought* 37 (2016) 46-63.

Lifschitz, Avi, 'The Arbitrariness of the Linguistic Sign: Variations on an Enlightenment Theme', *Journal of the History of Ideas* 73 (2012) p.537-557

Lifschitz, Avi, *Language and Enlightenment: The Berlin Debates of the Eighteenth Century* (Oxford University Press: Oxford 2012).

Lloyd, S.A., *Morality in the Philosophy of Thomas Hobbes* (Cambridge University Press: Cambridge 2009).

Locke, *An Essay concerning Human Understanding* ed. by Peter Nidditch (Oxford University Press: Oxford 1975)

Locke, *Ethica A* in: Mark Goldie (ed.), *Locke: Political Essays* (Cambridge University Press: Cambridge 1997).

Locke, *The Reasonableness of Christianity* in: *Works* VII (London: 1823).

Locke, *Two Treatises of Government* ed. by Ian Shapiro (Yale University Press; New Haven 2003).

Losonsky, Michael, 'Language, Meaning, and Mind in Locke's Essay', in: Lex Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 286-312.

Losonsky, Michael, 'Locke on Meaning and Signification', in: G.A.J. Rogers (ed.), *Locke's Philosophy, Content and Context* (Clarendon Press: Oxford 1994) 123-142.

Losonsky, Michael, *Enlightenment and Action from Descartes to Kant* (Cambridge University Press: Cambridge 2001).

Losonsky, Michael, *Linguistic Turns in Modern Philosophy* (Cambridge University Press: Cambridge 2006).

MacIntyre, Alisdair, *A Short History of Ethics* (Routledge: Oxford 2002 [1967]).

MacIntyre, Alisdair, *After Virtue: A Study in Moral Theory* (Bloomsbury: London 2011 [1981]).

Mackie, J.L., *Hume's Moral Theory* (Routledge: London 1980).

Mackie, J.L., *Problems from Locke* (Oxford University Press: Oxford 1976).

Magri, Tito, 'Locke, Suspension of Desire and the Remote Good', *British Journal for the History of Philosophy* 8 (2000) 55-70.

Malcolm, Noel, 'Hobbes's Science of Politics and his Theory of Science', in: idem, *Aspects of Hobbes* 146-155.

Malcolm, Noel, 'Hobbes and Spinoza', in: idem, *Aspects of Hobbes* (Clarendon Press: Oxford 2002) 27-52.

Mandeville, *An Enquiry into the Origin of Moral Virtue* in: idem, *The Fable of the Bees* ed. by Philip Harth (Penguin: London 1989).

Marshall, Terence, 'Epistemology and Political Perception in the Case of Rousseau', in: Eve Grace and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge University Press: Cambridge 2013) 76-122.

Martinich, A.P., *Hobbes* (Routledge: New York and London 2005).

Martinich, Aloysius, *Thomas Hobbes* (MacMillan: London 1997).

Masters, Roger, *The Political Philosophy of Rousseau* (Princeton University Press: Princeton 1968).

McNeilly, F.S., *The Anatomy of the Leviathan* (MacMillan: New York 1968).

Mcnive Hine, Ellen*, A Critical Study of Condillac's Traité des systèmes* (Martinus Nijhoff: The Hague 1979).

Melzer, Arthur, *The Natural Goodness of Man: On the System of Rousseau's Thought* (University of Chicago Press: Chicago 1990).

Mill, John Stuart, 'Bentham', *London and Westminster Review* August 1838.

Mishra, Pankaj, 'How Rousseau predicted Trump', *The New Yorker* (August 1, 2016).

Murphy, Mark, 'Desire and Ethics in Hobbes's Leviathan: A Response to Professor Deigh', *Journal for the History of Philosophy* 38 (2000) 259-268.

Nauta, Lodi, 'Hobbes the Pessimist?', *British Journal for the History of Philosophy* 10 (2002) 31-54.

Neuhouser, Frederick, *Rousseau's Theodicy of Self-Love: Evil, Rationality, and the Drive for Recognition* (Oxford University Press: Oxford 2011).

Oakley, Francis, 'Locke, Natural Law, and God – Again', *History of Political Thought* 18 624-651.

O'Neal, John, *The Authority of Experience: Sensationist Theory in the French Enlightenment* (Penn State University Press: University Park 1996).

Orain, Arnoud, 'The Moral Theory of Condillac: A Path toward Utilitarianism', *Revue de philosophie économique* 13 (2012) 93-117.

Pagannini, Gianni, 'Hobbes and the French Skeptics', in: John Christian Laursen and idem (eds.), *Skepticism and Political Thought in the Seventeenth and Eighteenth Centuries* (University of Toronto Press: Toronto 2015) 55-82.

Pagden, Anthony, *The Enlightenment and why it still Matters* (Princeton University Press: Princeton 2013).

Passmore, John, 'The Malleability of Man in Eighteenth-Century Thought', in: Earl Wasserman (ed.), *Aspects of the Eighteenth Century* (John Hopkins Press: Baltimore 1965).

Passmore, John, *The Perfectibility of Man* (Duckworth: London 1970).

Pécharman, Martine, 'Signification et langage dans l'Essai de Condillac', *Revue de Métaphysique et de Morale* 1 (1999) 81-103.

Perry, David, 'Locke on Mixed Modes, Relations, and Knowledge', *Journal of the History of Philosophy* 5 (1967) 219-235.

Pettit, Philip, *Made with Words: Hobbes on Language, Mind, and Politics* (Princeton University Press: Princeton 2008).

Pufendorf, *On the Duty of Man and Citizen* ed. by James Tully and trans. by Michael Silverthorne (Cambridge University Press: Cambridge 1991).

Putnam, Hilary, *Reason, Truth and History* (Cambridge University Press: Cambridge 1981).

Putnam, Hilary, *The Collapse of the Fact/Value Dichotomy and Other Essays* (Harvard University Press: Cambridge MA 2002).

Ricken, Ulrich, 'Linguistique et anthropologie chez Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 75-93

Rickless, Samuel, 'Locke's Polemic against Nativism', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 33-66.

Rieu, Alain-Marc, 'Le complexe nature – science – langage chez Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 27-46

Riley, Patrick, *Will and Political Legitimacy: A Critical Exposition of Social Contract Theory in Hobbes, Locke, Rousseau, Kant, and Hegel* (Harvard University Press: Cambridge MA 1982).

Rogers, G.A.J., 'The Intellectual Setting and the Aims of the Essay', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* (Cambridge University Press: Cambridge 2007) 7-32.

Roosevelt, Grace, 'An Alternative to Economic Man: The Limitation of Desire in Rousseau's Émile', in: Mark Blackell, John Duncan, Simon Kow (eds.), *Rousseau and Desire* (University of Toronto Press: Toronto 2009) 46-61.

Rorty, Richard, *Philosophy and the Mirror of Nature* (Princeton University Press: Princeton and Oxford 2009 [1979]).

Rosa, Rafaella de, 'Locke's Essay, Book 1: The question-begging status of the anti-nativist arguments', *Philosophy and Phenomenological Research* 59 (2004) 37-64.

Rosanvallon, Pierre, 'L'utilitarisme français et les ambiguités de la culture politique prérévolutionnaire (position d'un problème)' in: Keith Michael Baker (ed.), *The Political Culture of the Old Regime* I (Oxford 1987) 435-440.

Rosen, Frederick, *Classical Utilitarianism from Hume to Mill* (Routledge: London and New York 2003).

Rosenblatt, Helena, 'Rousseau, the "Traditionalist"', *Journal of the History of Ideas* 77 (2016) 627-635.

Rosenblatt, Helena, *Rousseau and Geneva: From the First Discourse to the Social Contract 1749-1762* (Cambridge: Cambridge University Press 1997).

Rossiter, Elliot, 'Hedonism and Natural Law in Locke's Moral Philosophy', *Journal of the History of Philosophy* 54 (2016) 203-225.

Rousseau, *Essay on the Origins of Languages* in: Gourevitch (ed.), *The Discourses and Other Political writings*.

Rousseau, *Émile* in: Gagnebin (ed.), *Œuvres complètes* IV.

Rousseau, *Lettre à Beaumont* in: Bernard Gagnebin (ed.), *Œuvres complètes de Jean-Jacques Rousseau* (Pleiade: Paris 1969) IV.

Rousseau, *Lettres écrites de la montagne* in: Gagnebin (ed.), *Œuvres complètes* III.

Rousseau, *Lettres Morales* III in: Gagnebin (ed.), *Œuvres complètes* IV.

Rousseau, *Notes sur "De l'esprit"* in: Gagnebin (ed.), *Œuvres complètes* IV.

Rousseau, *On the Social Contract* in: idem, *On the Social Contract, with Geneva Manuscript and Political Economy* ed. by Roger Masters, trans. by Judith Masters (Bedford/St. Martins: Boston 1978).

Rousseau, *On the Social Contract* in: Roger Masters (ed.) and Judith Masters (trans.), *On the Social Contract, with Geneva Manuscript and Political Economy* (Bedford/St. Martins: Boston 1978).

Rousseau, *Reveries* in: Gagnebin (ed.), *Œuvres complètes* I.

Rousseau, *Second Discourse* in: Victor Gourevitch (ed. and trans.), *The Discourses and other early political writings* (Cambridge University Press: Cambridge 1997).

Rousseau, Nicholas, *Connaissance et langage chez Condillac* (Droz: Geneva 1986).

Sacksteder, William, 'Hobbes's Science of Human Nature', *Hobbes Studies* 3 (1990) 35-53.

Sacksteder, William, 'Three Diverse Sciences in Hobbes: First Philosophy, Geometry, and Physics', *Review of Metaphysics* 45 (1992) 739-772.

Sangiacomo, Andrea, 'Locke and Spinoza on the epistemic and motivational weaknesses of reason: the Reasonableness of Christianity and the Theological-Political Treatise' *Intellectual History Review* 26 (2016) 477-495.

Schneewind, 'Pufendorf's Place in the History of Ethics', *Synthese* 72 (1987) 123-155.

Schneewind, 'Voluntarism and the Foundations of Ethics', in: idem, *Essays on the History of Moral Philosophy* (Oxford University Press: Oxford 2010) 202-221.

Schneewind, J.B., *The Invention of Autonomy: A History of Modern Moral Philosophy* (Cambridge University Press: Cambridge 1998).

Schosler, Jorn, 'Rousseau et Diderot, critiques de la philosophie égalitaire d'Helvéius' *Revue romane* 15 (1980) 68-83.

Schosler, Jorn, *John Locke et les Philosophes français. La critique des idées innées en France au dix-huitième siècle* (Voltaire Foundation: Oxford 1997).

Schouls, Peter A., *Reasoned Freedom: John Locke and Enlightenment* (Cornell University Press: Ithaca and London 1992).

Scott, John, 'Rousseau's Unease with Locke's Uneasiness', in: Eve Grace and Christopher Kelly (eds.), *The Challenge of Rousseau* (Cambridge University Press: Cambridge 2013) 295-311.

Scott, John, 'The Theodicy of the Second Discourse: The Pure State of Nature and Rousseau's Political Thought', *The American Political Science Review* 86 (1992) 696-711.

Sergeant, John, *The Method to Science* (London 1696).

Shaver, Robert, 'Grotius on Scepticism and Self-Interest', *Archiv für Geschichte der Philosophie* 78 (1996) 27-47

Shennan, Stephen, 'Regional population collapse followed initial agricultural booms in mid-holocene Europe', *Nature Communications* 4 (2013).

Shklar, Judith, *Men and Citizens: A Study of Rousseau's Social Theory* (Cambridge University Press: Cambridge 1969).

Simmons, John, *The Lockean Theory of Rights* (Princeton University Press: Princeton 1994).

Skinner, Quentin, 'Meaning and Understanding in the History of Ideas', *History and Theory* 8 (1969) 3-53.

Skinner, Quentin, *Reason and Rhetoric in the Philosophy of Hobbes* (Cambridge University Press: Cambridge 1997).

Smith, D.W., *Helvétius, A Study in Persecution* (Clarendon Press: Oxford 1965).

Sonenscher, Michael, *Before the Deluge: Public Debt, Inequality, and the Intellectual Origins of the French Revolution* (Princeton University Press: Princeton 2007).

Sorell, Tom, *Hobbes* (Routledge: London and New York 1986).

Sorensen, Leonard, 'Rousseau's Fulfilment of the Natural Public Law Tradition and his Contribution to its Demise', *The European Legacy* 10 (2005) 439-454.

Spragens Jr., Thomas A., *The Politics of Motion: The World of Thomas Hobbes* (Croom Helm: London 1973).

Stam, James, 'Condillac's Epistemolinguistic Question', in: R.W. Rieber (ed.), *Psychology of Language and Thought* (Plenum Press: New York 1980) 77-90.

Strauss, Leo, *Natural Right and History* (University of Chicago Press: Chicago 1953).

Takesada, T., 'Imagination et Langage dans l'Essai sur l'origine des connaissances humaines de Condillac', in: Sgard (ed.), *Condillac et les problèmes du langage* 47-58.

Taylor, Charles, *The Language Animal: the Full Shape of the Human Linguistic Capacity* (Harvard University Press: Cambridge MA 2016).

Trachtenberg, Zev M., *Making Citizens: Rousseau's Political Theory of Culture* (Routledge: London and New York 1993).

Trénard, Louis, 'L'influence de Condillac sur l'enseignement', in: Sgard (ed.), *Condillac et les problèmes du langage* 145-164.

Tuck, 'Optics and Sceptics: the Philosophical Foundations of Hobbes's Political Thought', in Edmund Leites (ed.), *Conscience and Casuistry in Early Modern Europe* (Cambridge University Press: Cambridge 2002).

Tuck, Richard, 'Hobbes'ss moral philosophy', in: Tom Sorell (ed.),*The Cambridge Companion to Hobbes* (Cambridge University Press: Cambridge 2006) 175-207.

Tuck, Richard, 'The 'Modern' Theory of Natural Law', in: Anthony Pagden (ed.), *The Languages of Political Theory in Early-Modern Europe* (Cambridge University Press: Cambridge 1987) 99-119.

Tuckness, Alex, 'The Coherence of a Mind: John Locke and the Law of Nature', *Journal of the History of Philosophy* 37 (1999) 73-90.

van Apeldoorn, Laurens, 'Hobbes on the Scientific Study of the Human Mind', *Archiv für Geschichte der Philosophie* 97 (2015) 308-333.

van Apeldoorn, Laurens, 'Reconsidering Hobbes'ss Account of Practical Deliberation', *Hobbes Studies* 25 (2012) 143-165.

Voltaire, *Correspondance* ed. T. Besterman (Geneva 1962).

Voltaire, *Letters on England* ed. and trans. by Leonard Tancock (Penguin: London 1980).

Voltaire, *Poème sur la loi naturelle* ed. by Francis J. Crowley (University of California Press: Berkeley 1938).

Watkins, J.W.N., *Hobbes'ss System of Ideas: A Study in the Political Significance of Philosophical Theories* (London: Hutchinson 1965).

Welch, Cheryl*, Liberty and Utility: the French Ideologues and the Transformation of Liberalism* (Columbia University Press: New York 1984).

Westerman, Paulien, *The Disintegration of Natural Law Theory: Aquinas to Finnis* (Brill: Leiden 1998).

White, Morton, *The Philosophy of the American Revolution* (Oxford University Press: Oxford 1978).

Williams, David, *Rousseau's Platonic Enlightenment* (Pennsylvania University Press: University Park 2007).

Wilson, Arthur, *Diderot* (Oxford University Press: Oxford 1972).

Wilson, Catherine, 'The Moral Epistemology of Locke's Essay', in: Newman (ed.), *The Cambridge Companion to Locke's "Essay Concerning Human Understanding"* 381-405

Wojciechowska, Wanda, 'Le sensualisme de Condillac', *Revue Philosophique de la France et de l'Étranger* 158 (1968) 297-320.

Wolterstorff, Nicholas, *John Locke and the Ethics of Belief* (Cambridge University Press: Cambridge 1996).

Wootton, David, 'Helvétius, from radical Enlightenment to Revolution', *Political Theory* 28 (2000) 307-336.

Yolton, J.W., 'Locke's Unpublished Marginal Replies to John Sergeant', *Journal of the History of Ideas* 12 (1951) 528-559.

Yolton, John W., *Locke and the Compass of Human Understanding* (Cambridge University Press: Cambridge 1970).

Yolton, John, *Locke and the Way of Ideas* (Thoemmes Press: Bristol 1993 [1957]).

Zagorin, Perez, 'Hobbes without Grotius', *History of Political Thought* 21 (2000) 16-40.

Zagorin, Perez, 'Hobbes's Early Philosophical Development', *Journal of the History of Ideas* 54 (1993) 505-518.

Zagorin, Perez, *Hobbes and the Law of Nature* (Princeton University Press: Princeton 2009).

Zinaich Jr., Samuel, *John Locke's Moral Revolution: From Natural Law to Moral Relativism* (University Press of America: Lanham 2006).

Nederlandse Samenvatting

Gedurende de zeventiende en achttiende eeuw vond er in de filosofie een revolutie plaats die een verregaande invloed heeft gehad op ons wereldbeeld. Aan de ene kant bepleitten wetenschappers en filosofen nieuwe methodes, gebaseerd op waarneming en experimentatie, om kennis te verkrijgen van de wereld om ons heen. Aan de andere kant waren filosofen op zoek naar een universele moraal – een objectieve wet, inzichtelijk gemaakt door middel van de ratio, die beschrijft wat rechtvaardig is, onafhankelijk van omstandigheden en heersende morele conventies (ook wel het natuurrecht genoemd). Dit proefschrift gaat over de vraag hoe het denken over deze twee vraagstukken elkaar beïnvloed heeft. En meer specifiek over de vraag hoe nieuwe inzichten over de oorsprong van kennis (epistemologie in vakterminologie) ook de aanzet hebben gegeven voor een verandering in het denken over de moraal (ethiek).

Eén van de belangrijkste ontwikkelingen in de epistemologie en ethiek gedurende de zeventiende en achttiende eeuw uitte zich in een fundamentele verschuiving in het denken over de rol en functie van de ratio (of: rede) binnen de menselijke geest. Tot halverwege de zeventiende eeuw waren veel filosofen er van overtuigd dat ieder mens, in ieder geval in potentie, de beschikking zou hebben over een onafhankelijke faculteit van de rede. Indien consistent geraadpleegd, zou deze rationele faculteit obectieve kennis kunnen verschaffen van zowel de moraal als de wereld om ons heen. Natuurlijk zag men in dat de ratio verre van onfeilbaar was. Bovendien zouden mensen continu verleid worden door de passies, irrationele verlangens en angsten, om van het pad van de rede af te wijken. Desalniettemin dachten veel filosofen dat kennis van de moraal en van de wereld om ons heen reeds latent aanwezig zou zijn in de menselijke geest.

De filosofen die in dit proefschrift besproken worden behoren echter tot een denkstroming die zich fundamenteel afkeert van deze manier van denken: het empirisme. De Britse filosoof John Locke wordt vaak gezien als de aartsvader van deze denkstroming, maar dit proefschrift probeert te laten zien dat ook zijn directe voorganger Thomas Hobbes, en achttiende eeuwse denkers als Condillac, Helvétius, en Rousseau als aanhangers van het empirisme geïnterpreteerd kunnen worden. In tegenstelling tot hun voorgangers, en sommige van hun

tijdgenoten, die vaak rationalisten worden genoemd, zagen empiristen de ratio niet als een autonoom of aangeboren onderdeel van de geest, maar simpelweg als de activiteit van het vergelijken en verbinden van ideeën. Locke en zijn volgelingen stelden bovendien dat de inhoud van deze ideeën geheel afkomstig is uit zintuiglijke waarneming. Volgens empiristen is kennis daarom niet slechts een product van een aangeboren rationele faculteit. Ware kennis komt tot stand door het vergelijken en verbinden van ideeën die we via onze zintuigen hebben waargenomen. De opkomst van het empirisme in de filosofie was dan ook nauw verbonden met de popularisering van wetenschappelijke methoden waarin waarneming en experimentatie centraal stonden.

Hoewel het empirisme van Locke en zijn volgelingen een aantrekkelijk antwoord gaf op de vraag hoe mensen wetenschappelijke kennis kunnen vergaren van de wereld om ons heen, was niet onmiddelijk duidelijk hoe dezelfde methode kan leiden tot inzicht in de moraal. Ook bij morele vraagstukken zien empiristen redeneren als niets anders dan het verbinden en vergelijken van ideeën. Maar wat is dan de oorsprong van deze morele ideeën? Het is duidelijk dat morele concepten zoals “rechtvaardigheid,” “vrijheid,” of “natie-staat” niet simpelweg door middel van zintuiglijke waarneming worden verkregen. Integendeel, deze morele concepten zijn door mensen bedacht, en hebben geen bestaan buiten de menselijke geest. Maar als de ideeën die ten grondslag liggen aan onze morele oordelen door ons zelf bedacht zijn, hoe kunnen we dan weten of deze overeenkomen met de universele moraal van het natuurrecht?

Deze vraag vormt het startpunt van mijn proefschrift. In mijn proefschrift tracht ik daarom te reconstrueren hoe we volgens vijf vroegmoderne empiristische filosofen objectieve morele concepten zouden kunnen verkrijgen, om daarmee, met gebruik van de rede, tot universele morele oordelen te komen. Daarbij is gebleken dat hoewel deze vijf denkers een groot aantal fundamentele veronderstellingen delen, hun antwoorden op dit vraagstuk sterk uiteenlopen.

Eén van de belangrijkste van deze gedeelde veronderstellingen was het idee dat morele normen de functie hebben om ons te leiden richting een zo lang en gelukkig mogelijk leven. Bij de compositie van morele concepten zouden wij daarom dit doel voor ogen moeten houden. Deze veronderstelling leidde echter direct tot een aantal nieuwe vraagstukken. Allereerst rees de vraag wat mensen eigenlijk gelukkig maakt? En in het verlengde daarvan: hoe voorkom je dat het streven naar geluk van één persoon ten koste gaat van het welvaren van anderen? Om deze vragen te beantwoorden dient men echter te weten hoe

mensen in elkaar zitten, wat voor gedrag ze vertonen, en wat hen drijft. Het beantwoorden van deze vragen, en daarmee het beredeneren van een objectieve moraal, vereist daarom een studie van de menselijke natuur.

Een fundamenteel onderdeel van deze studie naar de menselijke natuur was een theorie die beschrijft wat ons motiveert, hoe mensen besluiten om iets te ondernemen of te overdenken. Alle vijf filosofen die in dit proefschrift besproken worden waren het er daarbij over eens dat al onze daden en gedachten uiteindelijk het gevolg zijn van bepaalde verlangens of angsten, die collectief ook wel de passies worden genoemd. Met dit standpunt weken zij af van veel voorgangers en tijdgenoten, die in navolging van o.a. Plato dachten dat de menselijke wil bepaald zou worden door ofwel de passies, ofwel door een autonome rationele faculteit. Zij ontwaarden in de menselijke geest een voortdurende strijd tussen enerzijds de passies, die ons zouden verleiden om ons over te geven aan kortzichtige verlangens en angsten, en de ratio, die ons inzicht zou geven in een juiste en rechtvaardige manier van leven. De empiristen die in dit proefschrift besproken worden hadden de menselijke geest echter juist ontdaan van deze autonome rationele faculteit. Zij meenden dientengevolge dat menselijke handelingen en gedachten juist altijd voortkomen uit de passies. De rol van de ratio zou daarom beperkt zijn tot het beredeneren van de meest effectieve manier van handelen om onze verlangens te realiseren.

Bij deze laatste stelling, die geldt als de standaardinterpretatie in de secundaire literatuur, plaatst dit proefschrift wel een kanttekening. Moderne interpretaties stellen doorgaans dat de vijf empiristen besproken in dit proefschrift de ratio geheel ondergeschikt maken aan de passies. Daardoor zou deze beperkt zijn tot het vinden van manieren om verlangens te realiseren, en daarmee slechts een instrumentele rol spelen in de wilsvorming. Dit proefschrift probeert echter te laten zien dat in ieder geval volgens de filosofie van Hobbes, Locke, en Rousseau, de ratio óók deelneemt aan de wilsvorming door de passies te beoordelen op basis van de consequenties die het najagen ervan zou kunnen hebben. De functie van de ratio is daarmee niet slechts instrumenteel. De ratio bepaalt ook welke zaken het werkelijk waard zijn om te verlangen of te vrezen.

Op basis van deze veronderstellingen over de menselijke natuur construeren de vijf filosofen die hier besproken worden hun theorie van de moraal. Het is met name op dit punt dat zij van mening verschillen – zowel over de juiste aanpak als in hun conclusies. In het eerste hoofdstuk bespreken we de morele epistemologie (= theorie die beschrijft hoe mensen inzicht krijgen in de

moraal) van Thomas Hobbes (1588 – 1679), die voornamelijk bekendheid geniet vanwege zijn gedachte-experiment van de natuurstaat en zijn radicale herformulering van het concept natuurrecht. In de secundaire literatuur is er echter onenigheid over de juiste interpretatie van de rol van de ratio in zijn morele epistemologie. Sommige auteurs stellen dat wanneer Hobbes stelt dat het natuurrecht inzichtelijk kan worden gemaakt door middel van de rede, hij doelt op een aangeboren rationele faculteit. Dit proefschrift laat echter zien dat niet alleen in het geval van de wetenschap, maar ook wat betreft de moraal, de ratio voor Hobbes niets anders is dan de aangeleerde vaardigheid (verder ontwikkeld met behulp van taal) van het combineren en vergelijken van ideeën. Een andere veel voorkomende interpretatie stelt dat volgens Hobbes de ratio, en daarmee het natuurrecht, slechts een prudentiele functie heeft door ons de meest effectieve manier te verschaffen om onze verlangens te realiseren. In lijn met de eerder genoemde kanttekening over de juiste interpretatie van de relatie tussen de ratio en de passies, stelt dit proefschrift dat in Hobbes' deductie van het natuurrecht de ratio niet slechts een instrumentele rol speelt, maar tevens aangeeft welke passies aanwezig in de menselijke natuur het verdienen nagevolgd of vermeden te worden. Hobbes' gedachte-experiment van de natuurstaat dient er daarom toe om te demonstreren dat sommige verlangens, zoals het verlangen naar macht en rijkdom, uiteindelijk schadelijke consequenties met zich meebrengen. Tegelijkertijd laat het zijn dat sommige andere passies, zoals de angst voor de dood, het verlangen naar kennis en het streven naar een voldaan leven, mensen juist aanzetten om een sociaal, en daarmee gelukkiger, leven te leiden.

In het tweede hoofdstuk bespreken we de filosofie van John Locke (1632-1704), die naast zijn verhandelingen over politiek en het natuurrecht voornamelijk invloedrijk is geweest door zijn omvangrijke tractaat over epistemologie (*An Essay Concerning Human Understanding*). In dit hoofdstuk ontdekken we echter verschillende breuklijnen tussen Locke's natuurrechtstheorie enerzijds, en zijn epistemologie anderzijds, die een consistente interpretatie van zijn filosofie bemoeilijken. In de secundaire literatuur menen sommige auteurs dat Locke's morele epistemologie dezelfde methodologie gebruikt als die van Hobbes. Locke's conceptie van natuurrecht zou daarom gebaseerd zijn op een empirische studie van de menselijke natuur, en daarbij aangeven hoe mensen het meest effectief hun verlangens kunnen bevredigen. In dit proefschrift zien we echter dat deze interpretatie niet houdbaar is. Allereerst wordt deze weersproken door een fundamentele aanname

in Locke's epistemologie. Locke maakt namelijk een strikt onderscheid tussen ideeën van objecten in de wereld om ons heen, die wij door middel van de waarneming hebben verkregen, en ideeën van morele concepten, die door mensen zelf zijn bedacht. Het is volgens hem niet toegestaan om deze twee typen ideeën door elkaar te gebruiken en onze morele concepten te modelleren op empirische objecten. In feite zegt Locke daarmee dat het niet juist is om conclusies over hoe de wereld *is*, te presenteren als argument voor hoe dingen *zouden moeten zijn*. In lijn met deze positie stelt Locke dan ook dat een conceptie van hoe de menselijke natuur *is*, niet gebruikt kan worden om te beargumenteren hoe mensen *zouden moeten zijn*. Volgens Locke moeten we in de ethiek de mens dan ook slechts beschouwen als een rationeel wezen dat streeft naar de bevrediging van uiteenlopende passies. Een tweede veelvoorkomende interpretatie neemt deze laatste stelling als uitgangspunt, en schrijft Locke de positie toe dat de functie van de moraal simpelweg bestaat uit het voorschrijven van de meest effectieve manieren om onze verlangens te realiseren. Locke wordt daarbij gepresenteerd als een voorloper van het utilitarisme, een denkstroming binnen de ethiek die streeft naar het grootste geluk voor de grootste aantallen. In ons hoofdstuk zien we echter dat deze interpretatie wordt gecompliceerd doordat Locke vasthoudt aan het idee dat het natuurrecht een wet is die door God is uitgevaardigd. Volgens Locke is het primaire doel van het natuurrecht niet ons geluk tijdens het leven op aarde, maar wordt onze gehoorzaamheid aan God's wet pas beloond (of bestraft) in het hiernamaals. Omdat er, naast de interpretatie van de Bijbel, geen eenduidige manier bestaat om God's plan voor de mensheid te achterhalen, blijft het enigszins onduidelijk hoe Locke denkt dat we de voorschriften van het natuurrecht zouden kunnen inzien.

In het derde hoofdstuk behandelen we het denken van Etienne Bonnot de Condillac (1714-1780), een Franse filosoof die zich sterk heeft laten beïnvloeden door de epistemologie van John Locke. De filosofie van Condillac is daarbij onderscheidend vanwege zijn conclusie dat onze gedachten in veel gevallen gevormd worden door de taal die we spreken. De oorsprong van deze theorie ligt reeds besloten in de epistemologie van zowel Hobbes en Locke, die beiden al een nauw verband zagen tussen woorden en ideeën. Sterker nog, zij stelden dat bepaalde type ideeën, waaronder abstracte ideeën, maar ook morele concepten, letterlijk pas denkbaar worden in relatie tot woorden. En passant gaf Hobbes hiermee ook een verklaring waarom mensen in staat zijn tot verschillende

vormen van abstract denken, terwijl dieren, die immers ook zintuigen en een rudimentair verstand bezitten, beperkt blijven tot instinctief handelen doordat zij het vermogen tot het leren van een taal ontberen. Condillac werkt deze theorieën verder uit, en merkt op dat wij als individu niet vrij zijn om de betekenis van onze woorden te kiezen, maar dat deze in belangrijke mate wordt bepaald door conventies van de taal die we spreken. Condillac concludeert vervolgens dat het daarom onvermijdelijk is dat onze gedachten worden beïnvloed door de linguistische conventies die de betekenis van onze woorden bepalen. Condillac formuleert hiermee als één van de eerste filosofen in de geschiedenis een theorie van linguistisch relativisme, die stelt dat ons denken in belangrijke mate wordt bepaald door de taal die we spreken. Voor Condillac is deze conclusie echter problematisch, omdat linguistische conventies gedurende de geschiedenis geleidelijk en ongereguleerd tot stand gekomen zijn. Condillac gelooft daarom dat veel van onze linguistische conventies een inaccurate afspiegeling zijn van de werkelijkheid. Condillac denkt bovendien dat de linguistische conventies die de betekenis van morele termen bepalen vol zitten met historische gegroeide vooroordelen. Om deze redenen stelt Condillac een grondige hervorming van de taal voor, waardoor linguistische conventies een zo accuraat mogelijke afspiegeling van de werkelijkheid zouden worden. De linguistische conventies voor morele terminologie zouden daarbij in overeenstemming kunnen worden gebracht met het natuurrecht. Helaas presenteert Condillac weinig concrete plannen om dit laatste doel te realiseren. Condillac's filosofie is dan ook niet primair belangwekkend om zijn ethiek, maar eerder vanwege zijn bespiegelingen over de relatie tussen taal en het denken die, zoals we zullen zien, van grote invloed zijn geweest voor Rousseau's theorie van de menselijke natuur.

Het vierde hoofdstuk van dit proefschrift behandelt het denken van Claude Adrien Helvétius (1715 – 1771) die, anders dan zijn tijdgenoot Condillac, juist voornamelijk in morele en politieke vraagstukken geïnteresseerd is. Sterker nog, Helvétius' filosofische beschouwingen over moraal en de menselijke natuur kunnen gelezen worden als primair een theoretische onderbouwing bij zijn oproep tot concrete sociale en politieke hervormingen. Helvétius presenteert een sterk gesimplificeerde vorm van empirisme, die geïnspireerd lijkt door het werk van Locke en Condillac, maar de nuance mist van zijn voorgangers. Aan de ene kant stelt Helvétius naar aanleiding van Locke's stelling dat al onze ideeën te herleiden zijn tot de waarneming dat de menselijke geest in feite niets anders is dan een verzameling ideeën. Helvétius concludeert hieruit vervolgens dat ons

verstand wordt gevormd door de omgeving die we waarnemen, en presenteert daarmee de mens als een wezen dat geheel wordt bepaald door externe krachten. Aan de andere kant herinterpreteert Helvétius de stelling van Hobbes en Locke dat al onze daden en gedachten uiteindelijk te herleiden zijn tot verlangens en angsten, en stelt dat mensen in het geheel worden gedomineerd door hun passies. Helvétius lijkt daarmee te ontkennen dat mensen de mogelijkheid hebben om door middel van de ratio hun verlangens en angsten te beoordelen. Als gevolg van deze twee simplificaties presenteert Helvétius de menselijke geest als niets meer dan een passieve verzameling ideeën die wordt voortgedreven door het streven naar genot en de angst voor pijn. Deze reductionistische conceptie van de menselijke natuur ligt vervolgens ten grondslag aan de morele filosofie van Helvétius. In het verlengde van zijn stelling dat ieder mens wordt gedreven door een verlangen naar genot en een angst voor pijn, concludeert Helvétius dat de functie van de moraal is om zoveel mogelijk individuen de kans te geven hun passies te bevredigen. Met andere woorden, zowel de moraal als de samenleving heeft tot doel om te zorgen voor het grootste geluk voor het grootste aantal personen. Helvétius komt daarmee ontegenzeggelijk tot een vroege formulering van een utilitaristische ethiek.

In het laatste hoofdstuk behandelen we de filosofie van Jean-Jacques Rousseau (1712 – 1778) in de context van het empirisme zoals we dat in eerdere hoofdstukken hebben bestudeerd. Doordat Rousseau voornamelijk geïnteresseerd is in morele en politieke vraagstukken, en niet zozeer in epistemologie, wordt hij vaak niet als empiristisch filosoof gelezen. In dit hoofdstuk laten we echter zien dat Rousseau's conceptie van de menselijke natuur als veranderlijk, die ten grondslag ligt aan zijn ethiek, in belangrijke mate is afgeleid van zijn empiristische voorgangers. Om te beginnen neemt Rousseau van hen het idee over dat mensen redeneren op basis van ideeën afkomstig uit de waarneming en gestructureerd door middel van woorden. Rousseau gaat daarbij mee in de theorie dat abstract en moreel denken pas mogelijk wordt na het leren van een taal. Van Condillac neemt Rousseau bovendien de suggestie over dat taal een historisch construct is, dat in de vroege geschiedenis is ontstaan en zich geleidelijk ontwikkeld heeft. Gezien de belangrijke rol van taal in de ontwikkeling van het menselijk verstand, concludeert Rousseau hieruit dat de menselijke natuur gedurende de geschiedenis aan verandering onderhevig is geweest. Volgens Rousseau was de "natuurlijke mens", aan het begin van de geschiedenis en voor de uitvinding van taal, een dom maar gelukkig wezen. Net als veel dieren

leefde deze natuurlijke mens solitair, in harmonie met zijn omgeving, en waren zijn gedachten en verlangens beperkt tot zijn meest essentiele behoeftes. Na de uitvinding van taal konden mensen echter niet alleen met elkaar communiceren, maar kregen zij ook de mogelijkheid om complexe ideeën te construeren. Enerzijds stelde deze mogelijkheid hen in staat om op een abstracter niveau na te denken over de wereld om hen heen, met uitvindingen als landbouw en metaalbewerking tot gevolg. Anderzijds stelde deze mensen in staat om morele concepten, zoals “bezit” en “autoriteit”, te construeren en begrijpen. Hoewel deze twee nieuwe vaardigheden van de vroege mens het ontstaan van samenlevingen mogelijk maakten, is Rousseau van mening dat dit niet per definitie positieve ontwikkelingen zijn geweest – en wel om twee redenen. Allereerst heeft het ontstaan van taal er niet alleen voor gezorgd dat we slimmer zijn geworden, waardoor we beter zijn in het vinden van manieren om onze verlangens te bevredigen. De ontwikkeling van het verstand heeft tegelijkertijd ook gezorgd voor een explosieve groei van de passies – simpelweg omdat we beter op de hoogte zijn van alles dat we zouden kunnen verlangen of vrezen. Anders dan de mens in de natuurstaat, wiens passies in harmonie waren met haar behoeften, wordt de moderne mens geplaagd door verlangens en angsten die hij nooit in het geheel zal kunnen bevredigen. Volgens Rousseau zal de moderne mens daardoor altijd gefrustreerd blijven, en nooit werkelijk geluk vinden. Ten tweede meent Rousseau dat de ontwikkeling van de moraal met behulp van taal gedurende de vroege geschiedenis van de mensheid heeft gezorgd voor oneerlijkheid en onderdrukking. Geïnspireerd door Condillac stelt Rousseau dat gedurende de geleidelijke en ongereguleerde ontwikkeling van de moraal, deze vervuld is geraakt met vooroordelen in het voordeel van de bovenklasse van de samenleving. Eén van Rousseau’s oplossingen voor deze problemen bestaat uit een uitgebreid opvoedingsprogramma waarbij kinderen zo veel mogelijk worden geïsoleerd van de schadelijke invloeden van moderne samenlevingen. In dit proefschrift presenteren we echter ook Rousseau politieke theorie zoals beschreven in *Het Sociaal Contract* als een plan om deze gebreken van moderne samenleving op te lossen. Door te zorgen voor morele gelijkheid en politieke vrijheid streeft Rousseau’s politieke theorie namelijk naar een recreatie van de omstandigheden die de mens in de natuurstaat gelukkig maakten.

Dankwoord

Het afmaken van dit proefschrift heeft meer voeten in de aarde gehad, en bovendien aanzienlijk meer tijd gekost, dan ik van te voren had gedacht. Alweer zo'n 8 jaar geleden begon ik als afgestudeerd historicus bij de Faculteit Wijsbegeerte van de RUG aan mijn onderzoeksproject. De overstap van geschiedenis naar filosofie had ik behoorlijk onderschat. Mede hierdoor was het proefschrift nog lang niet af nadat mijn contract afliep, en heb ik dit in mijn eigen tijd af moeten maken, wat niet altijd makkelijk was. Gelukkig heb ik familie, vrienden, en collega's die mij tijdens dit proces met morele steun dan wel inhoudelijke aanwijzingen geholpen hebben. In het bijzonder zou ik willen bedanken:

Mijn promotor Lodi Nauta, die altijd nauw bij mijn onderzoek betrokken is gebleven, ook toen duidelijk werd dat het een project van de lange adem zou gaan worden. Je ontelbare inhoudelijke suggesties hebben dit boek gemaakt wat het is. Maar zonder je blijvende geloof in mij was het mij niet gelukt om dit proefschrift af te maken.

Daarnaast mijn tweede promotor Annelien de Dijn, die later bij het project betrokken raakte, maar desondanks erg belangrijk is geweest om mijn onderzoek richting te geven. Dank voor het altijd waardevolle commentaar bij mijn stukken en je blijvende betrokkenheid bij het project.

Mijn oud-collega's aan de Faculteit Wijsbegeerte, die vaak waardevolle suggesties hadden gedurende de WIP en Via Moderna colloquia. Met name wil ik Martin Lenz, Han-Thomas Adriaenssen, Andrea Sangiacomo en Bianca Bosman bedanken voor het lezen en uitgebreid becommentariëren van mijn stukken.

Alle vrienden die mij over de jaren moreel gesteund hebben. In het bijzonder zou ik graag Saco van der Ploeg willen noemen, die niet alleen de rol van paranimf op zich heeft genomen, maar ook het initiatief heeft genomen om een "vrijgezellenweekend" naar Gent te organiseren voor mijn "huwelijk met de

wetenschap," met de bedoeling mij aan te sporen het proefschrift nou eindelijk eens af te maken.

Verder ook Bart Geelen, die er niet alleen ook in Gent bij was, maar ook de fantastische omslag van dit proefschrift heeft ontworpen.

Daarnaast in, willekeurige volgorde, vrienden die altijd sterk betrokken zijn geweest bij mijn poging tot promoveren: Jaap Verheijen, Jan Gerlach, Luca Braccioli, Luuk de Goede, Floris Oskam, Ilse van Nierop, Renée Scheepers, Yvonne Jongma, Sander van Hees.

Tenslotte mijn belangrijkste steun en toeverlaat: mijn ouders, Jos en Marja, en mijn twee broers, Maarten (tevens paranimf) en Steven. Zonder jullie aanmoediging en adviezen was het me niet gelukt!